DOCUMENT RESUME

ED 440 661 IR 057 843

TITLE	Collections, Content, and the Web. Conference (Chicago, Illinois, October 5-7, 1999).
INSTITUTION	Council on Library and Information Resources, Washington, DC. Digital Library Federation.
SPONS AGENCY	Institute of Museum and Library Services, Washington, DC.
ISBN	ISBN-1-887334-73-4
PUB DATE	2000-01-00
NOTE	79p.; Joint partnership with the Chicago Historical Society.
AVAILABLE FROM	Council on Library and Information Resources, 1755 Massachusetts Ave, NW, Suite 500, Washington, DC 20036 ($20). Tel: 202-939-4750; Fax: 202-939-4765; e-mail: info@clir.org; Web site: http://www.clir.org/diglib.
PUB TYPE	Collected Works - Proceedings (021)
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	Information Management; *Libraries; Library Collections; *Museums; Nonprint Media; Online Systems; Technological Advancement; *World Wide Web
IDENTIFIERS	Digital Collections; Digitizing; Technology Role

ABSTRACT

Thirty leaders of museums and libraries met at the Chicago Historical Society (October 5-7, 1999) to discuss common questions and concerns about digitization of collections and explore the ways that the World Wide Web is affecting their collection-based institutions. This report presents the papers, under the headings of "Technology," "Audience," and "Collections," that were prepared in advance of the meeting and summaries, in each section, of the discussions they provoked. Following an introductory section by Abby Smith, papers include: "Mainstreaming Digitization into the Mission of Cultural Repositories" (Anne R. Kenney); "If You Build It and They Come, Will They Come Back?" (Katherine P. Spiess and Spencer R. Crew); "Library Collections Online" (Abby Smith); and "Museum Collections Online" (Bernard Reilly). The report ends with a concluding discussion and section outlining next steps. Appendices include a list of conference participants and a summary of the report, with tables and figures. (AEF)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 440 661

Collections, Content, and the Web

January 2000

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

B.H. Leneyr

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

● This document has been reproduced as received from the person or organization originating it.

□ Minor changes have been made to improve reproduction quality.

● Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Council on Library and Information Resources

IR057843

2

BEST COPY AVAILABLE

Collections, Content, and the Web

January 2000

Council on Library and Information Resources
Washington, D.C.

ISBN 1-887334-73-4

Published by:

Council on Library and Information Resources
1755 Massachusetts Avenue, NW, Suite 500
Washington, DC 20036

Web site at http://www.clir.org

Additional copies are available for $20.00 from the above address. Orders must be prepaid, with checks made payable to the Council on Library and Information Resources.

The paper in this publication meets the minimum requirements of the American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials ANSI Z39.48-1984.

Copyright 2000 by the Council on Library and Information Resources. No part of this publication may be reproduced or transcribed in any form without permission of the publisher. Requests for reproduction should be submitted to the Director of Communications at the Council on Library and Information Resources.

Contents

PUTTING CULTURE ONLINE 1

TECHNOLOGY

Mainstreaming Digitization into the Mission of Cultural Repositories, by Anne R. Kenney 4

Responses and Discussion 18

AUDIENCE

If You Build it and They Come, Will They Come Back?, by Katherine P. Spiess and Spencer R. Crew 22

Responses and Discussion 30

COLLECTIONS

Library Collections Online, by Abby Smith 34

Museum Collections Online, by Bernard Reilly 40

Responses and Discussion 48

CONCLUDING DISCUSSION 52

NEXT STEPS 55

APPENDIX I: Conference Participants 58

APPENDIX II: Assessing Institutional Web Sites, Summary of Report 60

Putting Culture Online

by Abby Smith, Council on Library and Information Resources

The World Wide Web is the brainchild of a consortium of academics who wished to create a content-neutral medium, open to all as a means of communication. It did not take long for the Web to be colonized by the commercial sector and, even more quickly, by a host of self-publishers posting materials of varying value, reliability, taste, purpose, and quality. As more and more information went up on the Web, public figures began to call for "quality content" on the Web, that is, things that have educational value and are created and maintained by trusted, brand-name institutions. Museums and libraries started receiving large sums from federal agencies and foundations, as well as digging deep into their own pockets, to digitize their collections.

How do museum and library collections translate into content on the Web? When art and research objects go from real to virtual, how does the relationship between object and viewer/user change? And who are the users of museum and library Web sites?

Thirty leaders of museums and libraries met at the Chicago Historical Society October 5-7, 1999, to discuss these questions and explore the ways that the World Wide Web is affecting their collection-based institutions. *Collections, Content, and the Web* was organized by the Council on Library and Information Resources (CLIR) and the Chicago Historical Society (CHS) and funded by the Institute for Museum and Library Services (IMLS). For many who came, it was their first opportunity to discuss common questions and concerns with peers from other cultural communities. Libraries and museums share few professional organizations, funding agencies, or external structures that regularly bring them together for substantive purposes. We took as our starting point one well-defined common feature of our institutions—the fact that we have been doing business (in some cases for more than 200 years) by collecting physical things in order

to make recorded knowledge and aesthetic experience accessible to our patrons. We chose to focus on three key issues—collections, audience, and, inevitably, technology. Because we asked questions about the relationship between collections and audience, we commissioned a survey of institutional Web sites to gather preliminary data about how sites have been conceived and for whom, and about who actually uses them.

Libraries and museums come to the Web with very different experiences of information technology. Libraries have long used automation for managing the description, cataloging, and inventory control of collections. They had used the Internet, the backbone of communication on which the Web ships its information cargo, long before the Web was created. This does not mean that they were necessarily early adopters of the Web, any more than were museums, which as a rule do not have the same robust technological infrastructures as libraries for management of their collections. On the other hand, museums in the last several decades have made great strides in making their collections more accessible to a large public and have developed intellectual, aesthetic, and educational portals for onsite visitors to their institutions.

Over the course of two days, participants at the Chicago meeting not only shared experience and expertise but also created a framework for an ongoing conversation that all hope will continue as we find our way in the new Web environment. The differences that became apparent between the operating assumptions of library and museum leaders were in some cases quite predictable. Perspectives on intellectual property, for example, diverged because of the traditional roles that libraries have played in the administration of fair use in the print world and the particular interest that museums have had in protecting the rights of those artists whom they display. Museums dealt forthrightly with issues of selection and presentation because they have a mandate to interpret. Librarians approached the matter of selection in some cases as if it were synonymous with censorship, because they traditionally place a high value on making information accessible without mediation. But in some cases the differences between types of museums (art or historical) and types of libraries (academic or public) were even more striking. In summarizing the discussions, we have tried to represent distinctly these four points of view—public and academic libraries, art and historical museums—to highlight the often-surprising intersections of values and concerns and the equally unexpected divergences of interest or experience.

This report presents the papers that were prepared in advance of the meeting and summaries of the discussions they provoked. It also includes the Web survey that CLIR commissioned from the Institute for Learning Innovation, which was designed to gather preliminary data about museum and library Web site design and use. There is no way that this report can capture the full flavor or content of the conversations that were begun in Chicago, but we hope that it serves to share many of the insights that participants brought to bear on a variety of topics. Most of all, this report attempts to present the frame-

work in which we hope to continue the conversations so fruitfully begun in Chicago.

We thank the Institute for Museum and Library Services for its support of the conference. The grant was part of its new effort to forge working partnerships between libraries and museums. We thank our partner, the Chicago Historical Society, which helped shape the program and created a hospitable atmosphere for our deliberations. We are especially grateful to those who came and gave their time and attention to the questions we posed. Their willingness to engage new and often difficult questions with candor and curiosity transformed our conjecture—that museums and libraries that digitize their collections have a lot to talk about—into a spirited and inspiriting exchange. Finally, we hope, through this report, to engage others who identify with the concerns aired here and wish to create collaborative structures for putting culturally significant materials on the Web.

Technology

Mainstreaming Digitization into the Mission of Cultural Repositories

by Anne R. Kenney, Cornell University Library

This conference on Collections, Content, and the Web brings together leaders from the museum and library communities to consider how the Web has affected the way we go about fulfilling our cultural mission. In this paper, I will address four topics that relate this technology to institutional responsibility, opportunity, and cost. My underlying argument is that cultural institutions face a point of critical transition. Over the past decade, they have come to appreciate the value of digital efforts to extend their reach. They must now appreciate that digitization is a normal part of doing business—one that is worthy of commanding its share of institutional resources.

Digital Collections Are Institutional Assets

As a normal part of doing business, institutions must create and manage their digital collections properly to ensure their long-term value and utility and to protect the investment that has been made in them. Although no universally endorsed guidelines or standards have been established for digital conversion of cultural resources, there is a growing belief in the value of creating "digital masters" that are rich enough to be useful over time in the most cost-effective manner. This position presumes that conversion requirements will be set at levels that are higher than either what is necessary to meet immediate needs or what is capable of being used under current technical environments. Michael Lesk and others have noted the economics of converting once (or, at least, only once a generation) and producing a sufficiently high-level image to avoid the expense of reconverting at a later date when technological advances either require or can effectively use a richer digital file (Lesk 1990). This economic justification is particularly compelling given that the labor costs associated

with identifying, preparing, inspecting, and indexing digital information far exceed scanning costs.

Institutional investments in creating high-quality digital masters are rewarded in the area of access and use. The library and museum communities are expressing a growing desire to develop cultural heritage resources that not only offer the broadest-possible use but also are comparable and interoperable across disciplines, user groups, and institutional types (NINCH 1999). Adopting a consistent approach facilitates integration between collections of images that artists and photographers are creating in digital form (the "born digital") and the "born-again" digital files that institutions create from their retrospective holdings. Peter Galassi, chief curator of photography at the Museum of Modern Art (MoMA), suggests creating a high-end digital master that is "purpose blind" (Sullivan 1998). Once created, the archival master can then be used to create derivatives to meet a variety of current and future users' needs. The quality, utility, and expense of various derivatives (e.g., for publication, image display, computer processing) will be directly affected by the quality of the initial scan.

In addition to the arguments for the economic advantages of converting once and for the creation of purpose-blind masters, preservation is the third main argument that is advanced for investing in rich digital masters. Digital files can be created to replace or reduce the use of deteriorating or vulnerable originals if the digital surrogates offer accurate and trusted representations.

But we do not decrease the preservation problem by relying on digital information; we only increase it. As Terry Kuny put it (1988), "Being digital means being ephemeral." Digital files must be created in a consistent and well-documented manner to make them worthy candidates for long-term retention. Disposition decisions should be based on continuing value and functionality, not limited by technical decisions that were made at the point of conversion or anywhere else along the digitization chain. We must appreciate how decisions that are made at the point of capture can affect our ability to manage, preserve, and use our digital collections.

Some guiding principles for safeguarding institutional assets include the following:

- Invest in the selection and creation of digital resources that have a high probability of use and reuse over time.
- Address preservation concerns from the ground up, including adequate quality capture and review; requisite metadata; and the use of standard, well-supported technologies. Unless these issues are addressed at the point of creation, "There is little prospect of archiving image resources that will survive technological change." (Ester 1996; see also Day 1998; NISO, CLIR, RLG 1999)[1]

[1] Day's work focuses on requisite preservation metadata. The NISO/CLIR/RLG initiative on standardizing metadata should provide specific preservation guidelines for digital image collections.

- Do not risk the master files by applying short-term solutions to short-term problems (many of today's constraints will not be tomorrow's, and we should avoid building an approach that becomes quickly outdated or superseded).
- Establish a social security fund for digital files from institutional resources (digital assets must receive perpetual care, which requires ongoing resource commitment).

Digital Collections Increase Patron Use, Which Places New Demands on Cultural Repositories

Cultural institutions experience incredible responses to digital resources that dwarf the use of their physical counterparts. The New York Public Library reports 10 million online hits a month, as opposed to the 50,000 books served at 42nd Street, and the Library of Congress transmitted nearly 347 million files in the first eight months of 1999 (Darnton 1999). These raw figures are not indicative of the qualitative use of this material; nonetheless, the ability to extend exponentially access to resources is compelling, particularly when developed for a museum, where a very small percentage of the total collection is ever on view at any one time.

Increased use is a double-edged sword, however, placing inordinate demands on resources of all kinds. Simply accommodating so many users requires institutions to support extremely powerful access systems. Peter Hirtle has noted the experience of the Church of Jesus Christ of Latter-day Saints, which in 1999 announced free access to many of its genealogical databases. Demand far exceeded expectations. The site had been built to handle 25 million hits a day—five times the anticipated use level. But in the first few weeks after it was opened to the public, the site recorded at least 40 million hits a day, and another estimated 60 million hits a day were turned away (Church of Jesus Christ of Latter-day Saints 1999).

A growing (and demanding) secondary clientele can tax staff resources. At Cornell, the Making of America Web site, consisting of 19th-century journals and monographs, receives 4,000 hits a day. A large share of the users is made up of non-Cornellians, who expect the digital library to act just like a regular library, replete with basic services. As the system becomes more stable, user requests have less to do with system difficulties and more to do with content inquiries, which often represent the interests of a general, rather than a scholarly, audience. Such questions as "What are my 1890s *Harper's* magazines worth?" make us feel a little more like an auction site than an educational site. Cornell began its digital library a decade ago under the rubric "any time, any place," and today must address the question of "anybody?"

The issues raised by user response to digital collections lead to the last two points I want to address: overcoming barriers to Web use and financing the enterprise.

Institutions Must Overcome Technical Barriers to Effective Use on the Web

Various user studies have concluded that all researchers expect the following things from displayed digital images:

- fast retrieval,
- acceptable quality, and
- added functionality.

Of course, they want many other things, too, such as the ability to print, to manipulate and annotate images, and to compare and contrast images. Increasingly, they want specialized services. In providing digital access, conflicts inevitably arise regarding what a user may want, what is affordable, and what the technology can deliver.

These expectations and inherent conflicts lead cultural institutions to confront a host of technical issues associated with quality, delivery, and utility that do not exist in the analog world. Unfortunately, no systematic assessment has been conducted to determine the cumulative effects of the total range of technological choices on the transmission and display of digital image material. File formats, compression processes, scripting routines, transfer protocols, Web browsers, processing capabilities, and the like combine to affect user satisfaction. This is particularly true when we consider the lag in technology adoption at the user's end. Users may think they want the highest quality, but they may be frustrated by how long it takes to download a file or may be disappointed when a beautiful color image displays in a largely posterized form.

Speed of Delivery

Speed of delivery is perhaps the major concern to users. A one-megabyte file might be accessed in a tenth of a second on a fiber network link but will take nearly three minutes on a v.90 modem. Because network configurations cannot be controlled, cultural institutions have focused on constraining image file size to speed access. Typically, institutions have reduced file size by limiting the resolution, or bit depth, or by applying compression. Each of these choices can have a pronounced effect on image quality. New and emerging file formats and highly efficient compression schemes such as Flashpix, GridPix, and Wavelet compression are gaining in popularity. They enable the delivery of large images over slow network links with little quality loss and offer the user the means to pan and zoom.[2] Another option for increasing delivery speed is to bundle images together, which may not increase the initial delivery speed but can facilitate "flipping" through a cache of downloaded images in rapid time. The most notable example of this capability is found in the use of Adobe System's PDF (portable document format) to view and print multipage documents. Other options include the use of multi-image TIFF

[2] Institutions experimenting with new file formats and compression schemes include the Library of Congress, the Library of Virginia, the University of Michigan, the U.S. Geological Survey, the Fine Arts Museums of San Francisco and the University of California at Berkeley, and the Cornell Johnson Art Museum.

(tagged image file format) files, CPC (Cartesian perceptual compression), and QuickTime movies.

The rush to embrace these new technologies should be tempered by the need to protect digital assets from obsolescence. This concern has sparked a continuing debate within the cultural community over the use of compression in master image files or the adoption of proprietary formats. As John Price-Wilkin has noted, "The Internet is littered with 'good ideas,' particularly in the form of impressive plug-ins or helper applications with frighteningly short life spans." (in press; see also Dale 1999)

The need to reduce file size to speed delivery may be a limited-term concern as broad bandwidth information pipelines and wireless high-speed data transfer capabilities are developed in the next 5–10 years to support research, electronic commerce, and entertainment. For instance, current Federal Communications Commission (FCC) rules require all analog broadcasts to be phased out by the end of 2006. The potential of digital television, in particular high-definition television (HDTV), to provide new and different kinds of information to a broad range of users—including access to digitized cultural resources—is tantalizing (FCC 1998). Beginning with Internet2, the U.S. government is funding efforts to build the Next Generation Internet (NGI) which will link research labs and universities to high-speed networks that are 100 to 1,000 times faster than the current Internet. Designed to handle high volumes of information, the NGI will make access to digital image files very easy and access to high-quality audio and moving-image transfer very practical (Cohen 1999).

Image Quality

Users expect digital images to offer visual quality comparable to that of the original. However, as has been noted, image quality may be reduced by the need for timely delivery. Quality can be further compromised by inadequate display technologies. Because monitor resolutions are often lower than those used to create digital image files, readers may be presented with difficult choices. They can choose a complete image, which can be delivered quickly but may be illegible; or they can examine image details but at the price of slow delivery and the ability to view only a fraction of the image at any given time. Color appearance is most problematic. The use of different browsers, the transfer between color spaces, or the reliance on underpowered monitors may affect it. Possible solutions include the use of sophisticated file formats such as portable network graphics (PNG), which supports both a Web-safe palette and sRGB, a color profile designed to ensure color consistency across platforms. Some institutions include gray-scale and color targets with their images to enable the end user to adjust the color when necessary. Others have created electronic targets and specified monitor settings to assist users in calibrating their monitors. Evidence suggests, however, that few users take advantage of these offerings. As Michael Ester has pointed out (1996), "The only controls that are apt to see widespread use are

those that are built into applications and underlying software." I suspect, however, that because color representation is a growing concern in electronic commerce, basic solutions will be forthcoming. As was learned in the mail-order business, no company can afford to handle too many returns and exchanges that are requested because the color of the ordered shirt does not match the color in the catalog—whether in print or on the Web.

Functionality

Digital image files are "dumb" files; they convey little beyond an electronic likeness of the original document or object. Additional work, which traditionally requires time-consuming descriptive cataloging or manual indexing, is needed to bring intelligence to these files. Containing costs while keeping pace with rising user expectations will require more automated image processing. Most of us are familiar with text conversion via optical character recognition (OCR) applications. These programs have improved tremendously, with error rates declining by half in the past few years because of advances in core recognition technologies, in weighted voting, and in the use of automated error-reduction applications. But highly accurate text conversion is still an elusive goal for most handwriting, for nonstandard scripts (such as Gothic), and for many nonroman languages (Dahl in press).

Interest in computer processing extends beyond textual information to graphic and photographic images. Raster-to-vector conversion software shows growing promise to create manipulable images for some graphic materials, such as maps, satellite and aerial photographs, architectural drawings, and engineering plans, but this capability still does poorly on rich, continuous-tone image files. Considerable research is under way in the area of content-based image retrieval (CBIR) to automatically extract features that characterize an image's appearance. Today's CBIR is based primarily on numerical measures of shape, color, and texture and is currently most effective where there is a need to retrieve information by image appearance (e.g., finding items of a particular color) rather than image semantics (e.g., pictures of children on a beach). Creative use of current capabilities can lead to retrieval either by characterizing the search in terms of proportion and color (e.g., a beach is 75 percent yellow, 25 percent blue) or by identifying a particular shape (e.g., a tiger), which will retrieve similar shapes and patterns that will include tigers but also fur coats. Because CBIR is actively being investigated, improvements could be rapid, but the capability to automatically retrieve images by a particular artist or photographs from a particular decade remains an elusive goal (Wu in press; Eakins in press; Lesk 1998).

In addition to providing added functionality, we can offer auxiliary features that facilitate more effective use of our collections. Consider, for instance, the success of Amazon.com, which is due in part to the added capabilities to facilitate access, selection, and ordering. Our digitized resources will be more accessible to a broader commu-

nity if we provide simple online tools that extend the capabilities of their analog counterparts, such as the following:[3]

- automated perpetual calendar, enabling a reader to key in month and day information (e.g., October 6) and receive a listing of all years in which that date falls on a particular day of the week (e.g., Tuesday),
- timelines to place historical items in the context of certain events,
- currency conversion tools that not only translate pesos into pounds but also peg value to their relative worth for any date in history,
- metric-to-English conversion tool,
- listing of scientific, medical, business, and cultural signs and symbols,
- multilingual dictionary and translation programs for text-searchable material,
- dimension tools not only to facilitate the use of digitized maps but also to enable the viewer to appreciate that a Dürer and a Dali may be of completely different scales (Handel 1995), and [4]
- lists of "sightings" in museum and auction catalogs.

Institutions Should Not Expect to Recover Costs Incurred in Digitization

No consensus has been reached about what it costs to create—much less maintain and make accessible—digital image files. The cost figures that are available vary tremendously, depending on the types of material being scanned, the image conversion requirements, the hardware and software used, and the range of functions covered in the calculations. There is no consistent price for outsourcing image conversion from vendor to vendor, or even from project to project, that is analogous to what we experience in other conversion efforts such as preservation microfilming.

We probably know the most about text scanning of disbound volumes, with estimates ranging from $.10 to $.30 per image for large production projects.[5] Figures for bound volume scanning are perhaps twice that amount. A number of institutions have found that they can obtain a better product and faster production rate when bound items are rendered into single leaves for scanning, even when the costs of rebinding are included (MacIntyre and Tanner 1998; ILEJ 1999).

[3] I am indebted to my archival colleagues at Cornell for many of these suggestions.

[4] Technical development at the Blake Archives Project includes a Java applet (The ImageSizer) to view Blake's work on screen at the actual physical dimensions. Available at http://www.iath.virginia.edu/blake/.

[5] These figures have been reported by Cornell, Michigan, and JSTOR (Journal Storage) (see also Odlyzko 1999). The Andrew W. Mellon Foundation has funded a project at the University of Michigan to document the full range of costs associated with digitization in a production environment. The results of that study will be available in late 2000.

Although production rates for film scanning theoretically are very high, in practice, current limitations pose difficulties that have reduced scanning rates considerably, and today, the costs of film scanning remain equal to or higher than the costs of paper scanning at the same resolution and bit depth. In a recent project to convert preservation-quality film, Cornell paid nearly twice what it would have paid a vendor to do single-sheet scanning. The Internet Library of Early Journals Project involved both bound-volume scanning and film scanning. The Project concluded that costs for microfilm scanning were higher and the quality lower than bound-volume scanning (ILEJ 1999).

Advances in grayscale capture will soon rival bitonal scanning in speed. When one moves from grayscale to color scanning, however, the time and costs increase significantly—on the order of two to three times.[6] Scanning figures for graphic materials represents an order-of-magnitude increase in cost over scanning text. Steve Puglia of the National Archives has completed a comparative analysis of digital imaging costs, the results of which have been presented in *RLG DigiNews* (1999). His findings offer a sobering reminder that imaging is not an inexpensive proposition. In the National Archives' Electronic Access Project, which included manuscripts as well as graphic and photographic materials, image acquisition costs averaged $7.60 per image. These figures go up when one considers high-end imaging projects of museum holdings; the reported production rates to create 70–100 MB files range from 15 to 70 images a day.[7]

More significant than image acquisition costs are the total costs associated with digital projects. Although figures vary from project to project, it appears that digital conversion represents one-third or less of total costs, with the other two-thirds going to metadata creation, administration, and the like. More sobering yet are the ongoing maintenance costs, which prove difficult to calculate because there are few production figures available. Some claim that the majority of costs are incurred in the first five years after creation and that they decrease significantly thereafter. Others claim that the maintenance costs will dwarf the costs of image acquisition. Unofficial figures from the Environmental Protection Agency peg the total costs of supplies, services, and hardware to maintain digital material for ten years at four to seven times the cost of creation. At a recent conference, presenters argued that digital images need to be migrated every three to five years at a cost equivalent to 50–100 percent of the costs associated with the original imaging project (Kenney 1997; MacTavish 1999).

[6] Gray-scale and color production figures reported in projects at the Library of Congress, the Smithsonian Institution, and at the Beinecke Library at Yale University.

[7] Production at the Johnson Art Museum at Cornell University averages 70 images a day for a nine-hour shift using one digital camera and two photographers. The Museum of Modern Art reports scanning and editing 20 images a day.

Most digital conversion projects have been funded by one-time appropriations from government, foundation, or institutional sources. Ultimately, an institution must assume the ongoing costs of maintaining its digital assets. Facing these costs leads to considering the economic sustainability of digital image conversion efforts. Such conversions can be accomplished in two ways. First, an institution can realize cost savings in other operational areas and divert those resources to the digital effort. This method seems more suited to libraries than to museums. Second, institutions can recover costs associated with digitization by selling or licensing their digital products. This approach may be more comfortable for museums than libraries, given their historic commitment to free access to information.

Reducing Institutional Expenses

The potential for cost savings is at the heart of one of the largest digitization projects today. JSTOR is based on the premise that space savings is a key cost factor for libraries, and these costs collectively will lead to the economic viability of digital conversion projects. JSTOR contends that a single library cannot save money by digitizing its older holdings, but that cooperative, multilibrary agreements might be economical. This cost assessment is based on the assumption that libraries can pay for their subscription fees by discarding paper holdings or by moving them to cheaper, less accessible offsite storage. By taking these actions, libraries presumably make space for other materials, reduce the need to build new libraries, and accrue additional operational savings (for example, in binding, preservation, retrieval, and reshelving). This model presumes that libraries can trust JSTOR to maintain its digital holdings in perpetuity. To date, the promise of cost savings has yet to be realized; few institutions, if any, have taken steps to wean themselves from the hard-copy versions. For the time being, one can assume that JSTOR members are subscribing to avail their constituencies of the enhancements and convenience it affords and thus have increased, rather than decreased, expenses. JSTOR offers little incentive to museums because they do not hold as many items in common and are even less likely than libraries to dispose of their physical collections (Guthrie 1999; De Gennaro 1997).

Recovering Costs

Some institutions hope to cover costs by generating revenue, which conflicts with the assumptions of many Web users that everything on the Internet should be free. Indeed, many institutions currently provide free access to their digital holdings, in part because they have received outside funds and in part because their administrations have supported the expense of maintaining the electronic presence. As institutions face the need to fund digital efforts from internal sources, the pressure to recover costs will grow.

A number of initiatives to develop cost-recovery solutions have been advanced, but little hard evidence is available to show that they will succeed. For example:

- In 1997, The British Library developed a business case to seek private sector collaboration to create a self-sustaining digital library service. Unfortunately, after a year, the library and the bidding consortium agreed to discontinue negotiations, maintaining that it had proved impossible to "balance the objectives of the Library with the commercial operating requirements of the consortium" (The British Library 1998).
- The University of Toronto has developed a business plan for selling paper versions of digitized books. Currently, its customers are limited to Japanese institutions, and the university is not breaking even. The university has concluded that the market must be expanded to make the program economically viable. The library has been willing to subsidize the operation to build its collection of digitized books but has not yet fully embraced this broader marketing strategy.[8]
- The MESL (Museum Educational Site Licensing) Project attempted to address many of the issues related to consortial licensing of museum images to universities for educational purposes. A detailed financial assessment concluded that consortial distribution of digitized museum objects to educational institutions will likely not be an economically sustaining, revenue-producing venture for some years to come. A collaborative initiative stemming from the MESL experience, The Museum Digital Library Collection, Inc. (MDLC), aspired to become a nationwide image licensing enterprise but now appears to be moribund (Besser and Yamashita 1998).[9]
- AMICO (Art Museum Image Consortium) is moving the concept of consortial licensing of museum images to educational institutions one step closer to large-scale reality. The project focuses on taking advantage of emerging education opportunities, but supporters also expect that it will bring new revenue sources and greater economic stability to the museums that participate in it (Bearman 1996; Trant and Bearman 1997). The Research Libraries Group (RLG) and AMICO have joined forces to provide access to the AMICO library. At this point, the effect of the Academic Image Cooperative, sponsored by the Digital Library Federation and the College Art Association and aimed at providing free access to art images for educational and nonprofit use, on AMICO's market is unclear.

This economic assessment leads one to question the cost-effectiveness of retrospectively digitizing library and museum holdings. Clearly, digitization efforts will have a greater chance of becoming sustainable if

- institutions consider digital material critical assets and create digital files in a manner to ensure their long-term value and utility;

[8] E-mail, Karen Turko to Anne R. Kenney, 31 May 1999.

[9] Information on MDLC is available at http://www.museumlicensing.org/.

- digital initiatives are mainstreamed within library and museum operations;
- libraries and museums can substitute digital for traditional means of access;
- researchers embrace the use of digital image collections and are willing to pay for some added value or convenience that digital versions offer; and,
- institutions are prepared to cooperate with one another to share the rewards and responsibilities of the digital world.

More important, cultural institutions should come to view digital conversion as a means to other things, not an end in itself. Susan Yoder, director of Integrated Information Services at RLG, has suggested that digitization efforts will be sustainable if they are justified by at least one other institutional goal beyond generating revenue (Yoder in press). For the foreseeable future, the digitization of retrospective collections will not pay for itself, but it may be a legitimate loss leader in a new service paradigm, enabling libraries and museums to compete successfully in reaching a broad range of cultural consumers.

REFERENCES

Web site addresses noted in this paper are valid as of January 20, 2000.

Bearman, David. 1996. New Economic Models for Administering Cultural Intellectual Property. Paper presented at the Digital Knowledge Conference, Toronto, Ontario, February 7, 1996. Also presented at EVA Florence, Italy, February 9, 1996. Available from www.archimuse.com/papers/db.mesl/economics.html.

Besser, Howard, and Robert Yamashita. 1998. The Cost of Digital Image Distribution: The Social and Economic Implications of the Production, Distribution, and Usage of Image Data. Report produced by the School of Information Management and Systems, UC Berkeley. Available from http://sunsite.Berkeley.edu/Imaging/Databases/1998mellon/.

Church of Jesus Christ of Latter-day Saints. 1999. New Family Tree Internet Service Swamped by Demand. Press release, May 26. Available from www.lds.org/med_inf/new_upd/19990526_FIGS_Demand.html.

Cohen, Elizabeth. 1999. Internet2 and the Next Generations. *NARAS Journal*. 8(2): 51–56.

Dahl, Kenn. In press. OCR Trends and Implications. In *Moving Theory into Practice, Digital Imaging for Libraries and Archives*, edited by Anne R. Kenney and Oya Y. Rieger. Mountain View, Calif.: Research Libraries Group.

Dale, Robin. 1999. Lossy or Lossless? File Compression Strategies Discussion at ALA. *RLG DigiNews* 3, no. 1 (February 15). Available from www.rlg.org/preserv/diginews/diginews3-1.html.

Darnton, Robert. 1999. The New Age of the Book. *The New York Review of Books*. March 18, pp. 5-7.

Day, Michael. 1998. Issues and Approaches to Preservation Metadata. Paper presented at the Joint RLG/NPO Conference on Guidelines for Digital Imaging, at the University of Warwick, Coventry, England, September 28-30, 1998. Available from www.rlg.org/preserv/joint/day.html.

De Gennaro, Richard. 1997. JSTOR: Building an Internet Accessible Digital Archive of Retrospective Journals. 63rd IFLA General Conference, Conference Programme and Proceedings, August 31-September 5, 1997. Available from http://ifla.inist.fr/IV/ifla63/63genr.htm.

Eakins, John. In press. Content-Based Image Retrieval. In *Moving Theory into Practice, Digital Imaging for Libraries and Archives*, edited by Anne R. Kenney and Oya Y. Rieger. Mountain View, Calif.: Research Libraries Group.

Ester Michael. 1996. *Digital Image Collections: Issues and Practice.* Washington, D.C.: Commission on Preservation and Access.

Federal Communications Commission (FCC). 1998 (November). Digital Television Consumer Information. Available from www.fcc.gov/Bureaus/Engineering_Technology/Factsheets/dtv9811.html.

Guthrie, Kevin M. 1999. JSTOR: The Development of a Cost-Driven, Value-Based Pricing Model. In *Technology and Scholarly Communication*, edited by Richard Ekman and Richard E. Quandt. Berkeley and Los Angeles: University of California Press.

Handel, Mark. 1995. Issues of Scale and Size in Visual Databases. Unpublished paper. Available from http://sunsite.Berkeley.EDU/Imaging/Databases/Fall95papers/handel.html.

Internet Library of Early Journals (ILEJ). 1999 (March). Final Report. Available from www.bodley.ox.ac.uk/ilej/papers/fr1999/fr1999.htm.

JSTOR. 1999. JSTOR: The Need. Available from www.jstor.org/about/need.html.

Kenney, Anne R. 1997. Digital to Microfilm Conversion: A Demonstration Project, 1994-1996. Final Report to the National Endowment for the Humanities. Available from www.library.cornell.edu/preservation/com/comfin.html.

Kuny, Terry. 1998. The Digital Dark Ages? Challenges in the Preservation of Electronic Information. *International Preservation News* 17 (May):13.

Lesk, Michael. 1990. Image Formats for Preservation and Access: A Report of the Technology Assessment Advisory Committee to the Commission on Preservation and Access. Washington, D.C.: Commission on Preservation and Access.

———1998. Finding Pictures. *RLG DigiNews* 2(1) (February 15). Available from www.rlg.org/preserv/diginews/diginews21.html.

MacIntyre, Ross, and Simon Tanner. 1998. *Nature*—A Prototype Digital Archive. Paper presented at the Sixth DELOS Workshop Preservation of Digital Information, Tomar, Portugal, June 17–19, 1998. Available from http://heds.herts.ac.uk/HEDCinfo/Papers/HEDSnature.pdf.

MacTavish, Sue. 1999. DoD-NARA Scanned Images Standards Conference. *RLG DigiNews* 3(2) (April 15). Available from www.rlg.org/preserv/diginews/diginews3-2.html.

National Information Standards Organization, Council on Library and Information Resources, and Research Libraries Group. 1999. Initiative on Standardizing Metadata. NISO/CLIR/RLG Technical Metadata Elements for Images Workshop. Available from www.niso.org/image.html.

National Initiative for a Networked Cultural Heritage (NINCH). 1999. Guide to Good Practice in Networking Cultural Heritage. Available from www.ninch.org/PROJECTS/practice/why.html.

Odlyzko, Andrew. 1999. The Economics of Electronic Journals. In *Technology and Scholarly Communication*, edited by Richard Ekman and Richard E. Quandt. Berkeley and Los Angeles: University of California Press.

Price-Wilkin, John. In press. Enhancing Access to Digital Image Collections: System Building and Image Processing. In *Moving Theory into Practice, Digital Imaging for Libraries and Archives*, edited by Anne R. Kenney and Oya Y. Rieger. Mountain View, Calif.: Research Libraries Group.

Puglia, Steven. 1999. The Costs of Digital Imaging Projects. *RLG DigiNews* 3(5) (October). Available from www.rlg.org/preserv/diginews/diginews3-5.html.

Sullivan, Terry. 1998. The Modern's Digital Archive. *Photo District News* (October/November).

The British Library. 1998. The British Library Digital Library Programme. Available from www.bl.uk/.

Trant, Jennifer, and David Bearman. 1997. The Art Museum Image Consortium: Licensing Museum Digital Documentation for Educational Use. Available from www.archimuse.com/papers/amico.spectra.9708.html.

Wu, Yecheng. In press. Raster, Vector, and Automated Raster-to-Vector Conversion. In *Moving Theory into Practice, Digital Imaging for Libraries and Archives*, edited by Anne R. Kenney and Oya Y. Rieger. Mountain View, Calif.: Research Libraries Group.

Yoder, Susan. In press. Sustainability Through Integration. In *Moving Theory into Practice, Digital Imaging for Libraries and Archives*, edited by Anne R. Kenney and Oya Y. Rieger. Mountain View, Calif.: Research Libraries Group.

Technology Responses and Discussion

Summary of Paper

Ms. Kenney summarized her paper and highlighted issues she wished to focus on in the discussion. She began by conceding that institutions are not ready for the transformation she called for—from working on digital projects to a mainstreaming of digitization programs. Too many institutions still view their digital files as surrogates, rather than as institutional assets, and are not prepared to support them as such. While funds for conversion may be easy to procure, funds to maintain these assets over time are nearly impossible to secure. This may be in part because of the uncertainties that characterize this period of transition: the cost savings that would argue for making digital services a core service of these institutions have not yet been realized; infrastructure problems have not been addressed; and staff are apprehensive about changes in jobs and usage patterns. Institutions should be focusing on value over the long term, especially the value that accrues to digital objects as more related materials from other institutions come online. Because of the nature of the Web, any one institution's digital collections are enhanced by—or compete with—those of others. Cultural institutions should now devote resources to maintaining digital resources on a secure financial, technical, and institutional base, and to engendering trust of digital formats among users.

Responses

Clifford Lynch, Coalition for Networked Information
José-Marie Griffiths, University of Michigan
Alan Newman, Art Institute of Chicago

There is a tendency to underestimate the importance of metadata costs, which can be half the total expenditure for a digitization project. Whereas we can reasonably expect the costs of conversion to go down over time as aspects of the processes are automated, we cannot expect the cost of creating metadata to decline. Moreover, we do not yet have a firm sense of what types of metadata are necessary. The creation of appropriate metadata depends on anticipating needs of the user. Do we really understand very much about our users? Do users want what libraries are offering—large collections of mainly unmediated and uninterpreted digital objects? Or do they want what museums traditionally offer—interpreted objects in a context that tells a story? We need to pay more attention to identifying our audiences and to developing technological responses suited to their needs.

Institutions should be preparing to deal with the larger implications of digital collections sustained over time. These include the responsibility to provide computer security and user authentication; the ongoing management of intellectual property rights in communities that comprise creators, publisher/distributors, and users; and the need to develop digital collections based on a business plan. A critical issue for museums and libraries is to determine what value is to be added through digitization and to base selection decisions upon that determination. This involves both a purposeful collection selection process for digitization and a reinforcement, in the finished product, of the crucial institutional brand recognition elements of reliability, quality, and trust. Given that libraries and museums are custodians of our culture's memory, it is their collective responsibility to grapple with these issues sooner rather than later.

The matter of selection—for whom and why—is crucial in developing a business plan, since every pixel carries a price. Another way of looking at the trade-offs that technology forces on institutions is to recall the classic triangulation formula—cost is quality times quantity—and remember the need to be practical when designing digitization programs. In theory, we should be responding to consumer demands, but how do we do that in (virtual) reality? Are color and scale standards suitable for the consumer market also sufficient for the more demanding work of art historians? If not, which standard should a museum choose? How can museums develop technical and intellectual guidelines that will anticipate the changes in scholarship and connoisseurship over time? Perhaps things such as image resolution and scale, which can be complex and expensive when aiming for the highest-possible quality, could be adjusted to suit the character of materials and their anticipated use. For example, paintings may dictate higher resolution capture than a numismatic collection. Cost-recovery schemes could reflect the same variations, with low-resolution images distributed free of charge and high-resolution images offered through a license fee. Finally, we should not let short-term concerns about delivery dominate thinking about how to digitize and at what level of quality. The pace of change at the delivery end, like so many other things, is beyond our control, but we

must pay attention to the long-term implications of our present-day choices. We should be developing guidelines for minimal levels of digital capture at the same time that we are monitoring the different type of access that our users are demanding.

Discussion of *Technology*

Costs

Academic librarians argued that cost recovery is becoming an imperative and that collaboration is touted as one way of cutting costs. Economies of scale provide one of the strongest motivations for collaboration. The librarians conceded that economies of scale have yet to be realized but asserted that the desire to avoid duplication is a strong factor in encouraging institutional cooperation. Looking at the example of off-site storage to control the costs of traditional book access, they proposed that a similar model could be used to control costs of storage and retrieval of digital data. Academic librarians also expressed concern about whether society will continue to transfer resources to digitization, especially since such efforts remain a cottage industry. There is not yet a mass market for the collections being offered; consequently, cost recovery is difficult.

Doubts linger about compelling incentives for collaborating. If information is increasingly seen as a commodity in the so-called knowledge economy, and if libraries are being urged to view digitized cultural property as an institutional asset, why should institutions collaborate? Moreover, technology has given users a sense of entitlement, even though there is no public recognition of the real costs of digital services. Large academic libraries that are used to collecting as comprehensively as they can in areas in which they declare an interest must decide whether to stop acquiring as many print items in order to digitize their existing collections and acquire new ones in electronic form only. Digitization is at risk of becoming yet another unfunded mandate.

Museum administrators were similarly divided on the issue of controlling the costs of digitization. Art museum managers were quick to claim AMICO as a signal success among art museums. With more than 50 members, AMICO provides a forum for museums to agree on standards for digitization and distribution of digital images, an achievement in and of itself. Art museum managers were also generally sanguine about the prospects for generating revenue and suggested user fees or graduated license fees for different levels of image quality as a way to recover costs.

History museum administrators were less optimistic that these approaches would prove satisfactory. They contended that collaborative databases such as AMICO and JSTOR are not generating new revenue, and that an assessment of the real costs of digitization must include the costs of not doing the things that institutions used to do. Revenue might have to be secured in a completely different way, such as taking advantage of the institution's reputation to sell advertising.

The digital environment is forcing both museums and libraries to think of themselves as businesses, something in which museums have more experience than libraries. (A perfect example of this dichotomy of models is the fact that the Chicago Historical Society charges a fee to visitors to exhibitions; access to its research collections, however, is free.) One promising solution to obviate the wasteful cottage-industry model that cultural institutions are all separately pursuing is collaborative marketing of resources and services. Museums may have more experience than most libraries with building market demand for their services, and libraries may have more understanding of the infrastructure needed to sustain digital programs. However, between the two types of institutions, the real costs of delivering digital products and services remain a daunting unknown. Perhaps most troublesome of all for mission-driven institutions is the recognition that museums and libraries must stop doing certain things in order to scale up their digital programs. There was a shared concern that these institutions cannot afford to let their missions compete with those of commercial endeavors such as publishers, entertainment conglomerates, online image archives, e-book services, and so forth.

Access and Control

Museums and libraries have traditionally been containers of information—able to control access to and use of their collections because of the constraints of time and place upon those who wished to use or study the physical objects contained in their collections. But museums and libraries do not contain their Web-accessible collections, and they cannot use traditional methods to control access to them or the ways in which they are used. People will come to these collections on the Web when and where they wish (though technology does allow access to sites be managed). In many ways, the Web is bringing museums and libraries closer together, at least in the minds of their virtual audiences, because their collections inhabit the same space now: cyberspace. In other ways, this meeting of library and museums is an extension of trends already present for other reasons. The recent experience of public libraries, especially those in major metropolitan areas, is that libraries are increasingly becoming museum-like; that is, they are becoming a central gathering place for people who are in search of a variety of activities and services. Museums are also places where people have experiences with objects that can be esthetic, emotional, or intellectual. The direct encounter with art cannot be replicated or replaced. Access to art differs from access to information; however, to the extent that both types of institutions are learning environments, both face the challenge of creating that environment online.

Audience

If You Build It and They Come, Will They Come Back?

by Katherine P. Spiess and Spencer R. Crew, National Museum of American History, Smithsonian Institution

The National Museum of American History (NMAH) is committed to inspiring a broader understanding of our nation and its many peoples. To accomplish this mission, we are exploring ways to further extend access to our collections and scholarship—access that has value and meaning to our audiences. The rapid development of the Internet and World Wide Web opens exciting opportunities to reach out to audiences across the country and around the world—not just to those who come to Washington, D.C. For those who do come to Washington or who visit our traveling exhibitions, technology offers the opportunity to see a greater number of objects and explore related topics more deeply while physically experiencing a few select artifacts.

The Opportunity and the Challenge

Objects in museums are objects out of context. The whole purpose of the modern museum, in terms of the public, is to create context—that is, meaning. In museums, we call the creation of context "interpretation." Multiple contexts can be created for and around a given object in our collection. Our central challenge in responding to the public through electronic media is how to address these varied contexts and nuances to serve a diffuse audience of varied backgrounds as well as multiple interests and needs.

The electronic media provide a way to look at a collecting arena in more depth: to create virtual exhibitions that address topics and themes not represented in our exhibition galleries; to present multiple points of view, interpretations, and experiences; and to respond to the needs of several different audiences. These media allow us to continue showing collections while one physical exhibition is being replaced with another and to explore various design and presenta-

tion approaches while experimenting with different script content. Technology also gives us a broader audience from which to gather evaluations and seek input during development—that is, before installing an exhibition in any of our galleries.

When museums exhibit objects in traditional ways, they must choose a few objects from their collection; for example, at any given time, less than five percent of NMAH's holdings are on display. By using a variety of multimedia and videoconferencing technologies for electronic outreach purposes, NMAH staff members provide access to more of our collections than we can ever make physically available at any one time. We create and deliver collection-based presentations to student and other groups, both in the museum and at remote sites worldwide. We develop electronic versions of the exhibits in our galleries and exhibits that exist only virtually. Our educators develop electronic interactive theme activities, curriculum materials, teacher guides, and teacher training materials for classroom and home use. And we have the opportunity to provide thematic access to our full collection of three million objects while referencing related materials held by other organizations.

Our goal is to have these products valued by educational institutions as well as by families and individuals. We hope that our products will increase understanding of the topics and issues that are presented and promote further exploration of them. How do we know when we have achieved our goal? A major challenge facing us is to develop methods for determining the value, meaning, and effectiveness of electronic products to our audiences. These audiences may be defined by age, ethnicity, educational background, physical and mental ability, and other variables.

Are the content and design of our electronic products useful and meaningful to the intended audience? Are we creating the intended learning opportunities? Is the content intellectually accessible, and are the words comprehensible to the intended audience? Just as important, does the audience find our site more useful and of more value than sites maintained by special-interest groups and for-profit organizations, who often present similar content in a more exciting and accessible way? These for-profit organizations are our real competitors. In the physical world, our audiences often pay more to use the products and services of these organizations than they pay to use ours. In the virtual world, the dollar cost tends to be the same—the cost of the Internet service provider.

We continue to hear about the number of "hits" that a home page receives, but knowing the number of hits does not tell us anything about the purpose or quality of the electronic visit. Online audience questionnaires tend to be completed only by those who like to express their opinions in this manner. Some museums have made use of focus groups, while others have partnered with teachers and school districts to develop and test specific Web products. There remains, however, a need to develop a complete and integrated evaluation process for museum World Wide Web offerings that reaches

out to all target audiences. NMAH is engaged in a search for such a process.

NMAH will develop a database that profiles our audiences and contains evaluative information about our electronic products. Collaboration among organizations that maintain this type of information benefits all. Technology provides tools to collaborate in developing and using this information, and we are interested in participating in such collaborations.

Our Expanding Audiences

The National Museum of American History is committed to being an audience-oriented organization. The full meaning of this commitment is still being discovered. Does audience-oriented mean that we provide our publics with what they want, what has meaning for them, what we believe they ought to have, or some combination of the above? How do we determine the answers to these questions? What criteria do we use to arrive at the answers? NMAH continues to struggle with these questions as it moves forward with its strategic plan.

In 1996, NMAH completed a yearlong survey to determine the characteristics and experiences of visitors to our building in Washington, D.C. As a result, we now know that a large proportion of our visitors travel to Washington from other parts of the United States (81 percent), primarily from the East Coast and the South (64 percent). Two-thirds of our visitors have completed at least a bachelor's degree. The average age of our visitors is 39 years, and 13 percent are members of minority groups. Visitors believed that NMAH's purpose is to engage in historical research and to educate. Overall, visitors favored public purposes (education, display, and entertainment) over professional purposes (history and preservation) by nearly two to one.

As part of its strategic plan, the National Museum of American History identified children and multigenerational groups as target audiences—onsite and online. We also stated our commitment to becoming a resource to those who live in the Washington, D.C., area. Over the next five years, our largest virtual audience will be children in grades five through twelve, teachers, and families. We base this determination on the fact that families and school-age children are two of our target audiences and that these groups are known to be frequent users of the Internet.

Although the academic community has always been both an audience and a collaborator, we also see a special relationship developing through the Internet between graduate programs in public history and the museum community. The goals of public history programs are different than those of traditional academic history programs. Museums can offer electronic access to scholarship and collections that help students learn about and better understand the issues of public history. NMAH is pursuing ways to develop electronic products that meet this need. Part of our approach will be to

form partnerships with academic programs in this area and with other museums.

We have defined our target audiences, but we also realize that the virtual visitor approaches the Internet with different expectations than those of a physical visitor who views an exhibition or attends a public performance. Users of the Internet come with specific questions or they come to "surf"—to see what they find. They expect a quick response to a question, and they expect to find enjoyment and discovery in surfing. Through links and surfing, some visitors find NMAH's Web site by accident; they do not necessarily realize that it is our site. How do we take advantage of these visits to create a repeat, intentional visitor?

The challenges we face in attracting Internet users and meeting their needs are no different than the challenges we face with visitors to our galleries. Electronic visitors, like visitors to our galleries, come with their own life experiences, vocabulary, thought processes, ways of learning, and agendas. As we explore different ways of using technology, how we address their expectations and characteristics will determine, to a large extent, the success of our efforts. As we reach out to diverse audiences, we must be sure that the themes and subjects meaningful to these audiences are in our products and in the databases supporting these products. The content must be intellectually and physically accessible to the target audience and the general public.

Although the challenges we face with visitors to both our galleries and our Web site appear to be the same, the environment in which these two groups interact with us is different. The virtual visitor has not come to the physical environment of our museum; we have been brought to his or her physical environment. In addition, the Internet allows that visitor easy and quick access to other sites that address the same or similar topics as those we present. Because of the technology, the viewpoints we put forth and the overall quality of our products are readily placed in a larger context and within a larger set of views. What makes our electronic presentations desirable and valuable in this context? Clearly, one way to find out is to involve our audiences in the development and evaluation of our electronic products. The question remains: How?

For our target audiences, NMAH is developing electronic products to advance family literacy, support and enhance school curricula, and provide families with resources to explore and learn about themselves as well as the history of this country and its peoples. We wish to make history fun and, among other approaches, will use electronic game formats for some of our products. In addition, we wish to explore creating an electronic interface that will enable visitors to generate their own exhibits through thematic access to our holdings and scholarship. If we are to be of value, the visiting public must be able to find us on the Internet, recognize us, and navigate our site. Therefore, we are planning a complete redesign of our Web site to increase our audiences and to increase the number of repeat visitors. Engaging our audiences in this process is a priority.

Approaches to Documenting Collections and Their Effect on Staff

Our desire to provide access to increasing numbers of our collections and to provide multiple points of view and interpretations brings with it the startling recognition that, all too often, extant documentation of collections, in both manual and automated systems, is woefully inadequate. The quantity of documentary information is generally more limited than one might expect; moreover, it represents only one viewpoint—that of the curator or specialist who brought the object into the Museum. There is no uniformity in the way information is organized. Neither consistency in the level or depth of detail noted nor conformity in vocabulary use or thesaurus controls can be found. In addition, the quality of recorded information often is either questioned by current staff or obviously inaccurate.

The root cause of much of the trouble is not hard to locate. Because of lack of staff and time, over the years, minimal information has generally been recorded about the items in our collections. For many items, we have only the object's name, the donor's name, a brief description, and an accession number. Museums require richer contextual information for research, interpretation, and exhibition activities. For example, information is needed on the history of the people and events associated with the object and on the object's origin, use, physical nature, and symbolic import.

As long as the Museum, with minimal amounts of staff, time, and expertise, continues to be responsible for managing new acquisitions and maintaining its existing collections, no opportunity for a critical review of the existing data and no possibility for analysis, upgrade, or revision of these data will arise. The situation thus becomes syndromic as continuation of the status quo only adds to the difficulty of resolving the problem in the future. How then do we correct and enhance the documentation so it can be used to provide better access to our collections and scholarship?

The textual content of a museum's collection documentation system—the nature and specificity of the terminology applied—depends entirely upon the training, work experience, and cultural background of the staff member who is documenting a particular item. Cultural diversity among staff offers the potential for rich documentation and interpretation of collections, and we must continue our efforts to build a truly diverse staff. During this era of staff reductions and tight budgets, however, we must find additional ways to bring this diversity to the documentation and interpretation of our collections.

Conversations among staff broaden viewpoints and enrich collections documentation. In the past, the content of these conversations was rarely recorded in museum cataloging systems. Such was the norm in most museums. Curatorial staff did not necessarily see the value in capturing and recording the information generated by these exchanges. Manual cataloging systems were not designed to hold multiple viewpoints.

Today, technology offers us an effective tool for capturing and acknowledging the contributions of multiple staff in documenting and interpreting museum collections. Before this tool is used for maximum benefit, curatorial staff must see the intellectual benefit to recording the full complement of information that is generated about an object. They must also see the value of using that information as they carry out research, create exhibitions, and develop public programs. They must believe that this work is as important and worthwhile as that involved in creating an exhibition. Collections documentation is an essential part of the foundation for our exhibitions and public programs, both physical and virtual. We must align our recognition and reward systems accordingly.

Our challenges now are to change existing norms and to create an environment where team, or collaborative, cataloging is an accepted part of processing new collections and of retrospectively documenting existing collections. More than that, we must develop a communal approach where expertise, experience, and cultural diversity are shared among cultural institutions and the academy, and where our audiences are invited to contribute to the documentation and interpretation of collections. Again, technology offers us the mechanism for this collaboration.

New Ways of Doing Business

The Internet not only facilitates collaboration in documenting and interpreting collections but also allows cooperative collecting and virtual lending. We can increase our collections—virtually—by adding to our documentation system information and images of objects held by others and that relate to our own holdings. As NMAH has developed Web products, we have inserted links to other sites with thematically related topics. We also would like to include objects—virtually—that are held by other institutions and individuals. Objects that are not available for physical lending may be lent virtually to enhance the electronic visitor's experience and increase the educational value of the product.

NMAH has formally moved to a team-based process for the development and presentation of physical and virtual exhibitions and of public programs. We recognize that the full range of expertise and knowledge needed to successfully complete these projects must be applied from the beginning of any project. The actual number of members on any team is dependent on how the expertise and knowledge are distributed among the staff. In some cases, one staff member may cover several areas; in other cases, one staff member may specialize in a single area.

All exhibits and public programs require content expertise—in-depth knowledge of the subject matter being presented. That expertise resides in the Museum's curators, historians, and subject specialists. Our products also require educators, collections managers for legal and ethical issues as well as preservation needs, and designers, producers, and public relations specialists. Designers and producers

who create physical exhibits and public programs need different skills than individuals who create Internet products. Recently, we have created a project management unit for our exhibition program. It is staffed by full-time professional project managers. The project manager is responsible for the timely completion of a project within budget. The project manager also is responsible for ensuring that concerns expressed by any member of the project team are addressed and that each team member has an equal voice. This structure frees the other team members from project management duties and allows them to focus on the work involved in their functional areas. In the past, most exhibition projects were managed by curators or subject specialists who, understandably, were often more sensitive to issues in their area of expertise than to the areas represented by other staff working on a project.

We see the Project Management Program as a way of bringing balance to planning and budgeting as well as of ensuring greater responsiveness to issues and concerns. We will apply this same management structure to our Internet products. Thus, we are about to issue our first formal electronic information management policy. This policy reinforces our belief that Internet products require the same critical review of content quality, the same consideration for audience, the same effort for design and production, and the same attention to evaluation as do our physical exhibits and public programs.

These efforts to better use the Internet and the World Wide Web are ambitious undertakings; however, we, like our sister institutions, have to move aggressively in this area. The visitors who enter our buildings now represent just a small part of our potential audience. Each of us must create a presence on the Internet, and we must be creative in the ways we position ourselves. It is not enough to simply announce our existence and the programs we produce. Information about our collections that includes images, interactive exhibitions, educational materials, and other engaging activities is essential.

High-quality Web sites are a resource-consuming proposition. These sites demand creative staff to develop them and competent staff to keep them up-to-date. Maintaining a Web site also puts pressure on staff to ensure that their collection information is current and that they integrate the Web site into their work and budgets as they create exhibitions. Moving in this direction demands a new mind-set. It demands that we recognize the growing and critical importance of the Internet and that we think about new ways to use this technology to our advantage. Not following this path may prove disabling or fatal.

Competition to use the information in our collections grows daily, particularly, among the for-profit sectors. Although we cannot expect to deny these sectors permission to use information in our collections, we need to provide our own products for the presentation of this information. These products must represent our philosophy about education and the presentation of our collections. We have important contributions to make, and we must let the public know what we have to offer. At NMAH, we are learning that this proposi-

tion is not a simple one. It can succeed only if we apply careful thought, creativity, and time and generate additional resources. Creating quality virtual products is a heavy burden for any museum to add to its commitments during already-challenging times. Nonetheless, work in this area must become part of the core mission activities of all of us in the future.

At NMAH, we are committed to building a presence in cyberspace that is engaging and that reflects our mission. In keeping with this commitment, our Web site must offer stories of our nation and its many peoples that have value and meaning to our audiences, must be based upon sound scholarship, must be created from quality designs, and must ensure intellectual and physical accessibility. If we accomplish these goals, we believe our electronic visitors will use us and, more important, will return on a regular basis.

Audience Responses and Discussion

Summary of Paper

Spencer Crew presented an oral summary of the paper he wrote with his colleague from the National Museum of American History (NMAH), Katherine Spiess. He emphasized that collection documentation and digitization need to accommodate a growing diversity of audiences and differing appreciations of objects. He called for a collaborative approach both within and among institutions for collecting, virtual lending, and creating links between Web sites. The technology almost mandates such collaboration, and new audiences demand it. The for-profit world is moving away from the concept of a captive audience; cultural institutions, which are in direct competition with entertainment conglomerates, need to become part of this environment of rapid change. What museums have uniquely to offer is the original, authentic object. Therefore, a museum site should offer the object in facsimile but also encourage audiences to experience the original object within the museum environment.

Responses

Laura Campbell, National Digital Library Program, Library of Congress
Samuel Sachs II, Frick Collection
Patterson Sims, Museum of Modern Art (MoMA)

Institutions need to create a digital presence in order to remain relevant and reach new audiences. When libraries approach these audiences, they need to provide the appropriate content, the necessary tools for navigation, and a high level of service. Achieving this is a process of trial and error. What does it mean for our institutions to become audience-oriented? Do we give people what they want or

what we think they should want? If we are in competition with consumer-oriented entities, we must distinguish ourselves by offering something unique. Libraries do not know their audience in the same ways that commercial interests know theirs, but they do know their collections. That is their competitive advantage.

For museums, the stakes may be different. There is a legitimate concern that in the race to be relevant, museums will be tempted to expand their audiences by compromising the high standards of display they maintain. For example, it is important not to oversimplify the content of a digital presentation of art in the mistaken notion that this will broaden its appeal. The personal experience between a person and a work of art is what museums are about, and the digital environment cannot provide that. For the Frick Collection, creation of a Web site to increase access to the collections does not mean, as it may in libraries, putting as much of its holdings online as possible. It means encouraging visits to the museum collections themselves. Encounters with original art are impossible in the digital environment. A comparison of two paintings by Constable, for example, would fail as a digital exhibition, because that can be done only by examining the originals side by side. Art museums must reach out to the public more than libraries in order to engage audiences and lure them into the physical environment of the gallery. This approach informs the Frick Collection's site, which offers a virtual tour of the museum as well as of parts of the collection.

There is an important distinction between exhibiting objects in a history museum such as NMAH and in an art museum such as MoMA. The NMAH may need, as Mr. Crew said, to create a context for its objects because they have been removed from the context that gives them meaning. The opposite is true for MoMA's obligations to its art objects. The real context for a work of art is the museum display. The Web cannot properly represent that object, no matter what context is created online, but it can provide rich information about objects and can do so round the clock, not only during museum hours. The Web can stimulate interest in the original object and increase audience size. The Web not only creates new audiences for art but also widens the circle of authorized interpreters. It expands authority for presentation from the traditional curator to the media or Web staff who create the digital version of an exhibition. The Web changes the balance of power in a museum. In the end, the media or Web staff for an exhibition will reach a wider audience than will the expert curator who relies on the traditional medium of the published catalogue. In contrast to libraries' claims that that they can satisfy the information needs of online patrons by serving digital surrogates of their holdings, a strong Web presence is not a surrogate for serving museum collections, which must be experienced firsthand. Digital initiatives at the Museum of Modern Art, far from reducing public interest in visiting the museum, have actually increased attendance to the highest levels in the institution's history.

Discussion of *Audience*

Administrators of art and historical museums dominated the discussion of digital audiences. This reflected in part the museums' recent focus on education and outreach, which are not traditional concerns of libraries. It also appeared to be a direct result of museums' understanding of the term "increasing access to collections" to mean bringing more people through their doors. This interest in attracting visitors was referred to as marketing or outreach, depending on one's point of view, but it was seldom called service in the librarian's customary use of the term.

Level of Presentation and Interpretation

The call for maintaining high standards in museums while reaching out to new audiences online resonated with many discussants. Some expressed concern that the number of visitors to museums may be soaring, but the intellectual and educational purposes of museums are in danger of being lost in the rush to marketing. Others said that the curatorial staff members of many museums voice the same concerns. Making intellectual content accessible to a broader audience is a professional challenge that curators must meet.

The Web makes possible a layering of access for different types of audiences, from those who are simply checking the hours a museum is open to those seeking to experience a virtual exhibition. How do we know who is looking for what? Institutions have insufficient data about how long visitors remain at a Web site or how often they bookmark a site or return to it later. There is very little meaningful information about how visitors use library and museum sites. The use of elapsed time of site visits as a way of measuring visitor interest is outmoded and irrelevant. Studies such as those at the Minnesota Historical Society indicate that there is no correlation between time spent by a visitor in a gallery and the extent of that visitor's appreciation or understanding of an exhibition, in part because visual learning is much faster than text-based learning. Again, the contrast between patterns of use for art museum Web sites and historical museum sites came into focus. Visitors to the latter are using them much as they would use library sites, that is, for personal research rather than in preparation for an exhibition experience. Both art and history museum administrators agreed that the Web makes it possible for museums to stay open 24 hours a day, and that ways should be found to make the content of Web sites available free to teachers and students. They also agreed that different levels of service and content can be provided to fee-paying audiences or to institutional or net members.

Librarians did not express the same interest in using digital media to create new audiences. They did note, however, that many of the issues that have dominated museums' concerns about audiences are now beginning to affect the physical environment of libraries. The architectural design of new public library buildings is being influenced by a desire to attract users to an environment that functions more like a museum. Even college and university libraries are now

investing resources to create learning environments within libraries, just as has been the case historically with museum exhibitions and programs.

Copyright and Intellectual Property

An overriding concern for all was the fact that copyright restricts much of what could be available on the Web. In some cases, the digital copyright laws are clear about what is permissible and what is not. Far too often, however, institutions are faced with ambiguous situations. When they are, they err on the side of caution and do not put things online. Once something is online, the institution loses control over its use. The user is always in charge. The idea of layering information for various users, some fee-based and some free, will not work in an environment in which institutions are constantly second-guessing themselves and cannot even decide, for example, if the use of thumbnail images infringes rights or invites lawsuits.

There were general assertions that museums tend to focus first on the rights of the creators, while libraries are concerned above all with fair use. This was related both to the differences in the ways in which libraries and museums are funded (as a rule, libraries do not charge admittance and thus have no financial incentive to encourage onsite visits) and to traditional notions of mission (museums have maximum control over terms of presentation, while libraries emphasize delivery of collections any time, anywhere). In practice, however, both types of institutions are behaving in similar ways. Both are wary of violating the rights of creators because they are valued donors. To alienate a creator is to risk warding off a future donor. The ambiguity of fair use and fear of litigation have kept both museums and libraries from moving aggressively in mounting multimedia collections, even if they are more amenable to digital presentation than traditional textual and fine arts visual resources. Finally, both museums and libraries are thinking about how to use what intellectual property rights they have to offset some of the costs of supporting digital services.

Collections

Library Collections Online

by Abby Smith, Council on Library and Information Resources

A cursory glance at a handful of library and museum Web sites reveals that these two kinds of institutions conceptualize their Web presences quite differently. Each of their core missions, though being in a broad sense about education, mandate different collections and, in turn, different service of those collections. Rather than discuss library and museum collections as a whole, therefore, Bernard Reilly and I will address them separately, even at the risk of emphasizing the distinctive rather than the common features of their Web sites. In the end, museum objects and library items are indistinguishable from one another when transferred to digital form. A digital *Blue Boy* and a digital *Huckleberry Finn* share the same behaviors, demand the same creation of metadata and management tools, reside on the same network, and are retrieved onto the same computer screen. Does this encourage virtual visitors to blur the distinctions between these entities in ways they could not if they were actually visiting a reading room or gallery? If so, does it make a difference?

The world of Huck Finn is one of ever-receding horizons and serendipitous finds. The general expectation of library visitors is that a library is like the Mississippi: all of life is there—whatever they want they can find—and it will cost them nothing to access but a little of their time and a very good river pilot. In reality, few library collections are as easy to navigate on the Web as they are onsite, but the exceptions appear to have learned a lot from museums and have focused on mounting highly curated objects of cultural interest rather than large, unedited collections.

Collections

Whereas art museums tend to collect the rare or unique, research li-

braries are more like natural history museums. They build collections that provide a comprehensive source base for researchers to use onsite. In many ways and, again, in contrast to art museums, libraries' collections must be redundant of other library collections in order to make things readily accessible to their patrons. Like natural history museums, they tend to have collection strengths in specific subjects, and, within those, frequently have scarce or unique materials of high artifactual value. Library special collections that contain unpublished and nonprint sources often contain items that are rare or unique. Taken item by item, they may be of minimal financial value and may have research value only to the extent that they are part of a larger whole.

Libraries value comprehensiveness because the only way to make a rich resource base available to researchers is to have as much onsite as possible. The availability of digital resources, with their promise of instantaneous access to information in remote locations, means that libraries do not have to own an item in order to serve it to their patrons. Redundancy of resources is not valuable in the digital library, where information is independent of any physical medium and access can be provided without proximity to the item. In fact, redundancy should be seen as a waste of resources. This limit on redundancy has serious implications for collection development—at least for collections of digital information—but it should also influence what analog materials are converted to digital form and how.

Services

The primary service that libraries offer is access to their collections—access to information about their collections and to the information contained in them. Unlike museums, libraries do not define access to include interpretation; on the contrary, the rawer the materials served, especially in a research library, the better. The more "cooked"—that is, selected, edited, shaped by an expert—the less integrity an item is deemed to have as an object of research. Librarians are not curators; they are not expected to have deep substantive knowledge about their collections, to provide historical and contextual interpretation, or to make judgments about objects. A subject specialist in a library should, rather, be expert in the source base of one or more domains of information to build an excellent collection that can be used and interpreted by the researcher. The librarian's responsibility is to acquire the best resources, organize them for ready access, and preserve them for future use.

What happens then when the real becomes virtual? How does digital transformation affect the library item, and how does that affect the services that libraries offer or, more precisely, the terms of service?

Digital Collections

Digitization of analog materials is, despite its revolutionary nature, usually treated as just a service—a service that provides new forms of access to analog materials. Digitization is essentially a superior form of copying—one without loss or, as we whimsically say, "lossless." It defies the physical constraints of time and distance and provides the chance to look at many disparate collections at a time and place convenient to the researcher. It can also significantly enhance information retrieval from small, damaged, or poorly preserved items through various types of image manipulation. Given the as yet unsolved problem of digital longevity, digitization is not accepted as a form of preservation reformatting, as microfilming is. Digitization serves preservation goals only to the extent that a digitized copy of an item can be served in lieu of an original and, hence, can reduce the stresses of physical handling—not an inconsiderable boon to rare book collections and photo archives. Nevertheless, there are cases where the availability of an item through a digital surrogate actually increases demand for an original because digitization is, it turns out, also a form of advertising.

Among the unintended consequences of digitization is the creation of additional collection items—that is, we now have more items in the collections to care for. A digitized item is a surrogate but more than a surrogate, a copy but more than a copy. Because it transforms an analog item into a digital one, digitization fundamentally alters the way that the original object holds and conveys meaning

A semiotician might say that a digital version of, for example, a first edition of William Blake's *Songs of Innocence* signifies in a different way from the original tome and, therefore, it signifies a different thing. The content of the message has changed because the medium has changed. Although this might strike some as arcane to the point of irrelevance, at both deep and superficial levels, the analog and digital forms of recording information are structurally and functionally radically different. Therefore, a digital copy is a radically different thing than the analog text or image or sound from which it was created. Given that the digital version is expensive to produce and also has a great deal of value added, the institution that is in the business of extending access through digital imaging should also be in the business of protecting and preserving its new, digital asset.

More than museums, libraries will view the digital collection item as one that should be or has been permanently accessioned into their collections. The Library of Congress does not plan to rotate items in and out of its National Digital Library. The items on the Digital Schomburg site at the New York Public Library constitute not an exhibition, but a collection. The site is an information resource that will, in all likelihood, not be deaccessioned. And, as Anne Kenney points out, the existence of digital back issues of scholarly journals on JSTOR has not led to deaccessioning the physical copies. Where is the cost saving?

Digital Services

Digitization changes the nature of service to patrons. Service of collections across the Internet is different from service in the reading room in ways that both enhance and undermine the traditional ways of conducting research with primary source materials. The Blake site at the University of Virginia[1] not only allows one to spend time with a virtual version of *Songs of Innocence* but also to compare several different versions of the same text held at different locations. Books that have been digitized to allow full text searching are transformed from a simple hard-copy text into a database that can be queried in ways that were formerly impossible. This easy access can be extraordinarily addictive, and it is no wonder that digital patrons are called "users."

Selection

Not all information resources in a library's collection are good candidates for digital conversion. On the contrary, only a rather small proportion of items has potential to be digitized effectively. The following list describes some real-life constraints.

- *Technology.* The creation of digital surrogates is ideal for items that are rare or fragile. It is also ideal for items that share similar provenance or that in some way are deeply connected but are housed at different locations. On the other hand, computers lead to a bias toward text over image and toward still image over moving image or sound. Oversized materials, such as architectural drawings, and materials in poor physical condition also are inimical to digital capture at this time.

- *Intellectual Property.* Because of copyright and its discontents, we find a bias toward selecting materials for digitization that are in the public domain. This bias is a great boon to Victorian and early modern materials, not to mention to fans of *Beowulf* or Thomas Jefferson. But the idea that a library's information resources can be adequately or meaningfully represented by a preponderance of public domain materials is senseless. Libraries are committed to continuing the traditional service that they have been providing to patrons in the actual world, where information is always fixed to a medium. This service has been free, and, in all but exceptional circumstances, materials that are both in and not in the public domain are equally accessible. This tradition is not practiced online. It is a cause for great concern among most professionals, and it should be a cause for equal concern among us all.

- *Resources.* Digitization costs a great deal of time and money. The presumption that libraries can afford to serve their collections

[1] Available from http://www.iath.virginia.edu/blake/.

digitally in addition to serving them onsite or through traditional interlibrary loan must be questioned. How can libraries afford the digital services that they wish to provide? We face a risk that, in order to subsidize the large financial outlays that digitization demands, libraries will be tempted to extract profits from digitization of their "museum" pieces, that is, items such as rare manuscripts or books that are of high artifactual value. Library collections are assets, but they seldom have been income-producing assets. Poor Huck, casting a sidelong glance at Blue Boy in the museum, sees that digitization just might offer the opportunity to cash in on his library's moiety of what has recently been dubbed "cultural heritage."

Some Consequences

The Marketing of Collections

Treating information resources as cultural heritage that can be marketed is a new idea for libraries. How would such marketing work? If a library can charge for some things, should it charge for everything, just to keep its franchise intact and protect the brand name? Consider, for example, a private university that has extraordinary holdings of rare books—books that are available to undergraduates for use in course work and that are deemed a great resource for scholars worldwide. If these texts were digitized, they would add a significant asset to the global digital library. But would they at the same time diminish, even if slightly, the allure of an undergraduate degree from this university by making what could only have been purchased with tuition—access to these books—available to everyone with an Internet connection? No one has quantified the relationship between a library's collections and services and the price of an undergraduate or graduate degree; however, few potential students who visit a college or university have not been shown the library and told of the richness of its collections and the advantages of access to these collections.

The Role of Specialists

The process of selecting items for digitization requires librarians to assume new and often uncomfortable roles. They are forced to play editor, selecting parts of large collections to put online rather than serving up the whole and letting the user select. They are also asked to censor items in a collection that, although appropriate for service in a reading room, are not appropriate to be broadcast across the Internet. Librarians do not like to censor research collections. The need to create metadata has the potential to turn catalogers into curators, for creating metadata involves creating a context that provides layers of information to facilitate retrieval and interpretations.

The Role of the Institution

One of the things that makes most library sites less appealing than museum sites is the paucity of interpretive materials. Digital visitors

want information that has been assimilated. The Library of Congress's American Memory site is successful, in part, because of the way it is curated. Collections that would take days to track down physically on Capitol Hill are brought together on the Library of Congress home page and grouped coherently and thematically. This presentation masks the fact that, in real life, these collections are not ordered, described, or served in relation to each other. Sites that reflect more closely the actual physical arrangement of collections, such as those of Yale or the University of Michigan, leave the online visitor, like the onsite visitor, to go from one library collection to another, without a central portal. This makes the visitor feel like he or she is traveling the Mississippi without a pilot. Museum sites, by contrast, direct visitors to their digital collections and provide more information. Moving from the wide, expansive, barely mapped world of Huck Finn to that of the Blue Boy, we find the surrounding landscape tamer, the green, open spaces well manicured, and the paths better trod.

Collections

Museum Collections Online

by Bernard Reilly, Chicago Historical Society

Museums have traditionally existed to acquire, preserve, interpret, and present works of art and artifacts. In recent years, some major institutions have emerged that have modeled new museum functions, for example, celebration (Cleveland's Rock and Roll Hall of Fame Museum), remembrance (the U.S. Holocaust Memorial Museum in Washington, D.C.), and even advocacy (the Museum of Tolerance at the Simon Wiesenthal Center in Los Angeles). In addition, there are many children's museums and science museums that have no permanent or core collections. Nonetheless, the care and presentation of collections still consume the lion's share of the energies and budgets of most American museums.

Because gallery and wall space are finite, most museums possess more items than they can present to the public in their own galleries, through loan exhibitions, or in print publications. Unlike libraries, which aspire to provide access to all of their holdings, museums customarily make their collections available on a selective basis. These limitations impose a regimen for the selection of works to be displayed and published that involves interpretation and judgment. Art museums and history museums usually present objects in an instructive or narrative framework. In choosing items to be presented under their aegis, museums routinely make decisions regarding the quality and importance of those items.

Service Models and the Curatorial Role

Museum curators acquire works, document and organize them for access, and present some of them in more or less interpretive settings. In acquiring and documenting works of art and artifacts, curators, like librarians, have traditionally followed what economists

term the "just in case" service model. That is, they build collections of long-term historical value, artistic value, or both and then compile and generate authoritative, often voluminous, data on individual works for use by scholars, specialists, and (increasingly) lawyers, who may or may not materialize on any given day on the museum's doorstep. The curator's presentation of such works to the museum's larger public is most often not on a library-like, patron-by-patron basis but is in the context of an exhibition or a publication that is aimed at a larger audience.

The long-term maintenance and administration of museum collections are governed by an elaborate framework of law, policy, and practice that has evolved over many years. That framework addresses issues of retention, disposal, accessibility, and management of the museum's collection assets.

Digital Collections and Services

Digital technology and the network capabilities of the World Wide Web vastly expand the possibilities for the presentation of museum collections. Although museums have long been able to circulate collection objects in traveling exhibitions and disseminate images of them in publications, slide sets, postcards, and films, digital technology offers a means of dissemination that is far less costly and, in some respects, superior. Once a digital master of an object is created, reproductions can be made virtually cost-free, transmitted over the Web almost instantaneously to an unlimited number of users, and made available on this worldwide network round the clock. They can be kept indefinitely, without being taxed as inventory (at least not yet) or running the risk of fading or pigment discoloration. Use in digital form protects the original object from theft or damage.

The information attached to the digital image can also be easily revised and upgraded without the traditional attendant press costs. Artifacts in digital form can also be enriched and illuminated by hyperlink juxtaposition with related museum objects, texts, data, spoken word, music, still and moving images, and a host of other ancillary resources. The experience and understanding of the original work can be enriched beyond what is possible in print through the addition of software that provides valuable functionality, such as the ability to rotate a portrait bust in pictorial space or reassemble the far-flung fragments of a Roman mosaic.[1]

In effect, these new capabilities alter the economics of museum collections administration. Digital technology enables museums to make their collections more freely available outside their walls. It also allows museums to behave more like libraries by providing access to more, if not all, of their holdings, albeit in surrogate form. By

[1] See, for example, the Whitney Museum of American Art's *The American Century* Web exhibition (1999), which incorporates audio files for a director's tour of the exhibition, film clips, and interactive features such as a time line and notebook for assembling the viewer's own tour.

removing the limitations of wall and gallery space, along with printing and binding budgets, digital media enable museums to offer what Abby Smith, in her paper, refers to as "a comprehensive source base for researchers."

With these advantages come burdens. Like libraries and archives, museums now must manage the thousands of digital surrogates they create in addition to the original pieces themselves. In the object-oriented world of museums, the digital surrogate carries considerably less weight as a stand-in for the original collection piece than digital files of manuscripts or maps in a library setting. Yet increasingly, the cost and value of such surrogates are compelling museums to treat these images as valuable, collection-like assets and to devote sizable resources to their control, preservation, and use.

Current Applications

To what extent have museums chosen to exploit these new capabilities? A look at current museum Web offerings suggests that although museums have embraced the interpretive and analytical capabilities of the new media, most have not exploited the potential these media offer to deliver larger amounts of collections content. In the virtual world, as in the analog domain, most museums provide selections rather than comprehensive collections. Roughly summarized, these selections consist of the following:

- masterpieces and other works chosen to illustrate the richness and range of an institution's permanent collection
- selected items from exhibitions that the museums have mounted, hosted, or both
- highlighted individual works, with educational, analytical, or other contextualizing commentary

Most often, the works that do appear are presented in thumbnail or less than half-screen format and at a relatively low screen resolution. More extensive museum holdings in "higher fidelity" tend to be presented within restricted environments that are open to limited audiences. These environments include password-protected university Web sites and consortia databases of works of art that are disseminated on a site-license or subscription basis. The Art Museum Image Consortium (AMICO), as one such consortium, presents digital images of works of art owned by member museums on a Web site that is licensed for educational uses.[2]

A few museums have declared their intention to place on the Web either their entire holdings or significant portions of their collections. For example, in January 2000, The Metropolitan Museum of Art mounted all of the approximately 2,000 paintings in its European

[2] For more information on the wide range of considerations and options for dissemination of museum images, see Zorich's *Introduction to Managing Digital Assets: Options for Cultural and Educational Organizations* (1999), Steiner's "Controlling Our Images: Museums and the Licensing of Imaging Products" (1992), and Bearman's "New Economic Models for Administering Cultural Intellectual Property" (1997).

Paintings Department on the Web. Similarly, the Fine Art Museums of San Francisco, in an "effort to behave more like a resource and less like a repository," have begun to mount high-resolution images of their entire permanent collection on their *Art Imagebase* Web site. As a condition of membership, each of AMICO's full members must agree to contribute digital images of 500 works a year to the consortium database until their entire collection is documented.

Museums tend to revise their Web sites often, removing and replacing artifacts and works to provide fresh content to casual and repeat visitors rather than retaining them as permanent features. Exhibitions featured on the Web are sometimes archived for a period but rarely are kept indefinitely.

Selection Factors

The following factors seem to be driving museums' decisions as to what collections and collection items they put on the Web and under what conditions they do so.

Intellectual Property

The potential for unlimited, unauthorized copying and distribution of images posted to the Web exposes museums to liability for infringing creators' and publishers' copyrights and subjects' publicity rights, as well as for committing other violations of intellectual property. These restrictions make Web dissemination of all works that are not clearly in the public domain a matter of risk management or subject to the painstaking process of obtaining item-by-item clearances. Museums have dealt with this problem in various ways. The San Diego Museum of Art has embargoed the display on its Web site of any works from its modern collection because of copyright restrictions.[3] Fear of infringement has caused other museums to restrict reproductions to small, low-resolution display images.

Cultural Sensitivities and Community Standards

The desire to avoid the risk of offending constituent groups can also influence selection. On the Web, museums reach far beyond their traditionally self-selected audiences to a more diverse and unfamiliar community of users. With this expansion of a museum's constituency, the likelihood that the content it offers will be objectionable to someone increases accordingly. Paintings can violate community standards of decency, historical cartoons can be offensive to particular ethnic groups, and objects with special religious or cultural significance can be considered inappropriate for display in the "secular" environment of the Web.[4]

[3] Available from http://www.sdmart.com/files/collection_modern.html.

[4] In developing its Web site (http://www.si.edu/nmai), the National Museum of the American Indian has had to scrupulously avoid reproduction of Native American grave goods and other items associated with the dead.

Technology

Because current digital technology is limited in terms of display size and resolution, certain kinds of museum objects lend themselves to Web dissemination better than others. (Anne Kenney's paper discusses these limitations in depth.) Original works on paper, such as architectural drawings, maps, and posters with substantial text, as well as costumes and other works that require scrutiny of detail for full appreciation show poorly on the Web. Likewise, three-dimensional museum objects, such as sculpture, armor, and architectural fragments, often depend for their full effect on characteristics such as mass and scale, which are not communicable in a digital image.

Funding

Because digitization projects are expensive, they often must be underwritten by special funding, and the interests of funders often contribute to determining the content. What are major corporate, government, and private funders supporting on the Web? On the one hand, they are supporting educational materials and tools for the K–12 audience; on the other hand, they are supporting technologically innovative research and development projects. Museums' own revenue-generating activities can also influence their selection decisions, skewing content toward their more "licensable" images or, conversely, limiting display resolution on those same pieces to prevent loss of potential income through unlicensed use.

Availability

A prerequisite for presenting museum collections on the Web or in any medium is the availability of complete, authoritative, curatorial documentation and consistent, high-quality image capture. Although digital technology offers an easy means of dissemination, the cataloging and systematic duplication of sizable collections of museum objects can require many years to accomplish. Hence, holdings that are already well documented and photographed normally rise to the top of the list of museums' candidates for the Web.[5]

Marketing

Most museum Web sites are devoted primarily to promoting museum visitation, visibility, and products. Often more prominent than collections on these sites is information about the museum: how to get there, current exhibitions and programs, what to see, and what one can buy. Esther Dyson (1998) has said that the most scarce and valuable commodity in the Web economy is the attention of consumers, and that organizations will have to add value and functionality to their content to build audiences.[6] Sometimes this richness of content will be at the expense of breadth of content.

[5] The Metropolitan Museum of Art acknowledged that the availability of comprehensive documentation and photography was an important factor in the museum's decision to mount its entire European paintings collection on the Web.

[6] The Minneapolis Institute of Art has used collection images as the basis for a value-added "product" by offering ready-made electronic postcards of works of art in their collections on its Web site.

Distribution of the Selection Decision

The complexity of the technology involved and the scale of the financial commitment required by digital technology tend to remove museum Web content decisions from curatorial departments, the traditional loci of museum exhibition and publication decisions. Where the responsibility and, often, the impetus for determining Web content reside vary from museum to museum. Some Web projects originate in technology or imaging services departments, some in marketing departments, and others the museum director's office. A number of museums, like the Art Institute of Chicago and the Museum of Modern Art, have set up Web editorial committees that have representation from both senior management and curatorial offices.[7] In other museums, such as the National Museum of the American Indian, the publications department administers the Web site relatively autonomously.

Future Digital Collections

Beyond the near term, what will library and museum Web offerings eventually add up to? For purposes of discussion, here are three of the many possible models or approaches that museums and libraries might adopt to deliver collections electronically. Although neither exhaustive nor mutually exclusive, these models represent fundamentally different strategies for populating the Web.

- *The Selective Model* presents selected works from library and museum collections. Because it is subject to the internal and external selection factors cited above, this model could result in significant blind spots in library and museum Web offerings. Notably absent from these offerings, for example, might be items of a potentially problematic or disturbing nature, works produced after 1923, works and artifacts considered less accessible to a broad audience, and poorly documented or unattributed works. Conceivably, the public could be left with low-resolution images of "safe" and popular materials from the precopyright period. These offerings, moreover, would be fluid rather than permanent.

- *The Collections Catalog Model* would provide Web delivery of high-quality digital versions of whole library and museum collections or cohesive bodies of materials from single libraries or museums. These would be accompanied by authoritative information about the works and artifacts to facilitate searching and analysis and, possibly, by software tools for manipulation and analysis. Delivering these materials is extremely costly in terms

[7] The Metropolitan Museum of Art's Museum Internet Task Force is chaired by the museum's senior vice president for external affairs and includes among its members both the director and president of the museum.

of preparation time, technology, and (potentially) rights clearances. The materials would also require long-term maintenance.

- *The Shared Holdings Model* is a modern variation on the traditional *catalogue raisonné*. This model involves assembling on the Web digital "collections" of like or related objects that are drawn from the holdings of several museums, libraries, or both. The shared holdings approach has been used successfully with holdings of ancient papyri, medieval manuscripts, and the like, where the multiple parts of an object or set of documents are held by different institutions. Like the collections catalog model, the shared holdings model is a very resource-intensive approach. Digital resources created through this model would also have to be maintained and updated indefinitely. Costs could be shared among participating institutions.

Concluding Thoughts

Although museums and libraries will continue to differ in the materials they collect and make available onsite, putting collections online poses many of the same constraints and challenges to both kinds of institutions. Given the extraordinary expense involved in creating and maintaining digital collections, the biggest challenge they both face is to determine which collections work online and which do not. When a painting or a book is digitized, what makes that individual artifact unique and uniquely interesting is lost, and what remains is a bitstream. By virtue of their traditional missions and resources, museums and libraries see the opportunities offered by online collections quite differently, and that is reflected in their presences on the Web. If, as Anne Kenney argues, libraries and museums are developing a new service paradigm, then we must study the behaviors not only of users online but also of the objects, or the digital artifacts, with which they interact.

REFERENCES

Web site addresses noted in this paper are valid as of January 20, 2000.

Bearman, David. 1997. New Economic Models for Administering Cultural Intellectual Property. In *The Wired Museum*, edited by Katherine Jones-Garmil. Washington, D.C.: American Association of Museums.

Dyson, Esther. 1998. *Release 2.1: A Design for Living in the Digital Age*. New York: Broadway Books.

Fine Art Museums of San Francisco 1998. *Art Imagebase* [Web site]. Available from www.thinker.org/imagebase/index.html/.

Steiner, Christine. 1992. Controlling Our Images: Museums and the Licensing of Imaging Products. *Museum News*, (July–August): 62–64.

Whitney Museum of American Art. 1998-1999. *The American Century* [Web exhibition]. Available from www.whitney.org.

Zorich, Diane M. 1999. *Introduction to Managing Digital Assets: Options for Cultural and Educational Organizations*. Los Angeles: Getty Information Institute.

Collections Responses and Discussion

Summary of Papers

In the oral summary of her paper, Ms. Smith emphasized the differences between digital content on the Web and traditional library materials. Digital objects have no analog in the analog world. Digitization for the Web is in many respects like publishing: it requires selection, editorial judgment, cataloging, and metadata. These decisions are presenting libraries with new questions. In the conventional reading room, no one was denied access to a book because it was under copyright, yet current copyright law makes it impossible for libraries to fulfill their traditional role as providers of access to all information. Mr. Reilly emphasized that technology is enabling museums to disseminate their collections more widely and to larger audiences. In this sense, museums are becoming more library-like. A new set of skills is being demanded of museum staffs in order to create and manage the digital environment; however, in the end, it is still a question whether the product is larger than the sum of its parts.

Responses

Christine Steiner, Esq.
Leila Kinney, Massachusetts Institute of Technology
Francis X. Blouin, Bentley Historical Library, University of Michigan

Copyright issues are eroding the traditional boundaries between creators, managers, and users. Much in the copyright regime that was possible in part due to the slow pace of creating and disseminating knowledge is gone for good in the fast-changing digital world. The court interpretations have shifted in favor of the copyright holders, and the number of materials in the public domain is shrinking.

Nevertheless, it is important to remember that the value-added aspects of library and museum work are copyright protected.

Scholars are hoping that the Web will make teaching and research easier and more productive. Why can't museums help overcome growing barriers to research and publishing in art by becoming more like libraries and making their collections available to a broader audience? Libraries present materials in ways that are good for pedagogy and scholarship, with as little interpretation as possible, so that users can find their own meanings. Museums have been loath to present their collections on the Web without full interpretation, yet they have not drawn on research by faculty and students who have expertise about museum objects. Doing so would increase the density and pedagogical depth of museum sites on the Web.

While museum and library collections may serve distinct purposes in the analog world, in the digital we are entering a slow, incremental process of transformation of all cultural resources. How can museums and libraries maintain the institutional context of the digitized items on their Web sites? Digital collections exist outside the confines of museum and library buildings, whose very architecture is evocative and creates a context for understanding the origins and uses of the items they house. How much of our current decisions about digitization are being shaped by past experience and the traditional organization of information? Must we continue to define content in terms of institutional boundaries and proprietary control? Will the time come when the Web is the institution and we are all artifacts of an age gone by?

Discussion of *Collections*

The complex and interrelated issues of selecting and editing collection content for digitization, and related concerns about intellectual property and rights holders, provoked the most sustained discussion of the conference.

Selection

In meeting the challenge to mount significant portions of their holdings, museums are stymied by the limited amount of information on the objects in their collections and by the difficulty of creating the documentation necessary to digitize these objects. Several approaches were suggested to accelerate the processing, cataloging, and creation of metadata for collections. For example, the model of engineering schools, in which teams of students are assigned to evaluate and describe a collection or group of materials as an academic project, could be adopted. While this approach seems particularly well suited to the interactive Web, the implications would be that museums would cede full responsibility for every identification or interpretation of an object. The trade-offs between increased access and authoritative information, however clear, remain difficult to resolve.

There was discussion about the degree to which museums and libraries should prepare their digital objects for consumption on the Web by creating contextual information. The assumption that libraries present materials on the Web that are uninterpreted or unmediated and that museums place objects in a context does not hold up under scrutiny. The very act of putting materials up on the Web constitutes selection and interpretation. Some warned that a naïve public might perceive the digital selection as a complete truth rather than an interpretation. History museum administrators suggested that what is needed on cultural sites is a clear identification of selection criteria and the underlying assumptions and approaches to the material. The Web does not provide the user a means of seeing the infrastructure of selection or description, or the original context or provenance of items. Some argued that, on the Web, the public wants and will accept multiple voices and interpretations as long as they are attributed to a source. As an analogy, the public understands that a filmmaker has a point of view and accepts this as a convention of story telling. Since institutions have no control over the use of digital content, they have no choice but to try to understand and react to the patterns of use of those visiting their Web sites and should bear the responsibility of making institutional choices clear.

Art museum administrators concentrated on how digital content is to be shaped within an institution. A focused dialogue between curators, the traditional controllers of interpretation, and professionals within education departments of museums should result in a greater diversity of views to which Web users may respond. The Web, with its democratizing technology that invites interaction and multiple viewpoints, is highlighting the tension between the curatorial and educational departments in many art museums. On the other hand, there is already a difference between the exhibition narrative and the exhibition catalog, in which the material is presented with much fuller interpretation and argument. The same differences in level of interpretation can be maintained on the Web, where the presentation of digital surrogates of objects can be accompanied by a variety of interpretations for which the author is clearly identified. The Web makes it possible for the museum to be a place where wider points of view are welcome without necessarily making the museum responsible for what is said.

While acknowledging that all scholarship depends on selection of the primary source to be preserved and made accessible in libraries and archives, academic librarians expressed great concern about any approach to digitization that would edit or bowdlerize the full content of the original collection. Scholars select items within a collection to mount and argue an interpretation of the facts, analogous to the way in which museum curators select materials to tell a story in an exhibition. Libraries need to recognize that in putting materials on the Web, they are creating a collection that others will use to tell stories that are different than the ones that scholars put together. Recent trends in scholarship and teaching show that visual resources are an increasingly important source of information. It is thus all the

more important that images be fully identified. Also essential are an explanation of the relationship the objects on the Web have to other objects in the same collection and a description of why they were chosen. As educational institutions have long understood, material cannot be presented without an appropriate context. In any event, the Web will always have limitations as place of scholarship because the nature of some research material renders them inappropriate for Web dissemination. Consider the example of a bomb recipe found in the Chicago Historical Society's Haymarket collection. An archivist argued that comprehensive digitization should mean including everything in a collection: items considered objectionable, such as that explosive recipe, should not be silently eliminated. But does such a recipe really belong on the Web? A history museum administrator countered that instructions for making a bomb do not belong in a story being told by that project, although they belong in the collection that contains them.

Intellectual Property

Art museum administrators maintained that their institutions are effectively managing intellectual property issues by recognizing the significance of copyright and engaging in direct negotiations with artists and copyright holders, even though these are sometimes difficult. They are also entering consortial agreements such as AMICO, which provides images for nonprofit use. Art museums can also control the distribution of digital images of their collections by using sur-prints within images or by adding value to a digital image and then melding the image and added context together to produce the originality required for copyright.

While conceding that art museums in general have been much more aggressive than history museums in dealing with intellectual property issues, history museum administrators expressed concern about how unprepared many museums are for addressing the problem. Many museums are very naïve about violation of copyright on the Web; they put up digital images for circulation without negotiating full rights for the material. Institutions that are conscientious about property rights are severely constrained in creating virtual exhibitions on the Web.

Librarians have traditionally been able to control access to their holdings and apprise users of their obligations to respect copyright requirements. The digital environment, however, is threatening these traditional barriers to improper access and use. Library databases can be restricted to authorized users. The real concern of libraries now is in determining what is in the public domain, since once a text is put up on the Web, it is completely out of the library's control. A search for rights holders can be time-consuming and fruitless, especially for performances.

In the end, digital copyright issues put museums and libraries in the same situation. Museums are in direct competition with the copyright holder or Web user, who can download images and build a site overnight. Consequently, they have no more control over their digitized collections on the Web than do libraries.

Concluding Discussion

During the plenary session that concluded the conference, each participant had the opportunity to define the principal issues that had been raised and propose the next steps that need to be taken.

Academic librarians stressed the need to consider taking steps on a much larger scale than before. These steps include building the infrastructure needed to sustain digital assets and finding the permanent funding to do this. They urged that we should be thinking about building common databases, not individual Web sites, and addressing the many issues around convergence. Amalgamation is the national business trend, and cultural institutions would do well to create a "mega" Web site--something like a "culture.org" or "content.edu"—that would create a digital presence of sufficient scale to match the power and scale of commercial endeavors on the Web. As the economy is increasingly being driven by commercial knowledge creation, e-companies are acting as both creators and publishers of digital content. Cultural institutions must build relationships with the technologists, businesspeople, and legislators who so deeply affect the fate of cultural institutions in the online world. Copyright is another important issue that must be addressed directly. Historically, academic institutions have tended to duck this matter because they have been able to control access to collections. While they can build walls around collections online in some cases, copyright still severely restricts what is mounted on the Web. There is little choice in the future but to engage both rights owners and users. Another pressing concern is the long-term maintenance of digital files.

Finally, academic librarians called for more pilot projects designed to develop best practices for collaboration, to collect data on the impact of the digital environment on resource acquisitions bud-

gets, to cost out the creation and preservation of electronic exhibitions, and to track life cycle costs of digital materials. This is a collaborative task and it will bear fruit only to the extent that we engage our primary users—students and faculty.

Public librarians agreed that a much larger scale Web presence is required—one that would include a huge array of integrated resources that could be accessed easily from a single portal. There is a need on the Web for a critical mass of digital material in support of wide-ranging research and real education for many varied audiences. Public librarians cautioned that ways must be found to make the Web available to all audiences, especially those children not well-served by the public education system. We must develop new models of service, not simply replicate of patterns of access that existed before the creation of the Web.

Art museum administrators urged broader efforts to integrate digital programs of cultural institutions with external sources of support. It is important to establish collaboration not only with important members of the business community, they argued, but also with the larger nonprofit sources of funding. As a model digitization program, art museum administrators expressed great confidence in AMICO; one administrator called it a model database and research tool, and another asserted that AMICO has already achieved much of what the conference participants had been discussing. From the administrators' perspective, the real problem is not generating the digital images but creating the documentation that must accompany them. This, and the resolution of continuing intellectual property issues, are the most important problems to be solved.

History museum administrators pointed out the ways in which the traditional boundaries between libraries and museums are being blurred, just as the historic balance of decision-making authority within museums is being eroded. There is also a blurring of the profit/non-profit distinction, with commercial companies becoming creators of content. The result is an inescapable "commodification" of knowledge. They also stressed that the traditional audience for cultural institutions is not only expanding, but is being transformed in such a way that it is no longer possible to identify a museum audience or a library audience. The audiences are blurring, and institutions cannot maintain their traditional position as gatekeepers who can control or select their audience. If there is one thing we have learned about the Web, it is that quality is not more important than, or even as important as, ease of access. Users decide how, when, and where they will use an institution's materials. In fact, most people do not even distinguish between a museum and a library site. They are looking for information. The academic users present underscored this notion, claiming that the important issue is not which institutions digitize which materials or on whose Web site they reside. The crucial need of academic users, both faculty and students, is to have the broadest possible access and the most powerful searching tools to locate digital resources wherever they may be found.

Costs are also a major concern. One history museum administrator reiterated his concern about the sustainability of the AMICO model, both as a continuing and reliable source of revenue and as an agency for the permanent maintenance of digital resources. History museum administrators joined their art museum colleagues, however, in stressing the need for new efforts to provide better descriptive control over collections. They also agreed that libraries need to follow the lead of museums in developing a more sophisticated attitude toward marketing and in addressing issues of intellectual property.

Next Steps

The fundamental challenge at the moment is to determine what steps will ensure that the Web can be greater than the sum of its parts, that is, that the museum and library presence on the Web amounts to more than a cluster of individual Web sites. No one believes that the Web will replace libraries and museums, but many can see a time when the Web blurs and eventually erodes, in the user's mind, the current distinctions between libraries and museums. The Web is allowing the creation of new spaces in a landscape different from that of either libraries or museums. We are rapidly moving into an environment in which preconceptions formed by traditional institutional associations and proprietary control are being challenged and dissolved. How do we facilitate the migration of culturally significant artifacts and information into the Web environment and make those sites as easy to use as the best commercial sites? How do we present our collections when the Web becomes the portal to the institution?

Participants called for the following actions:

Continue the dialogue and widen the participation

- The opportunity for museum and library professionals to exchange views informally and frankly should continue, perhaps under the auspices of CLIR.

- The number and type of institutions at future meetings should be expanded to include natural history, ethnological, or archaeological museums; a different mix of local and state historical societies; special collections; and more public libraries.

- Cultural institutions should establish areas of agreement and common purpose upon which to build an agenda and seek external funding.

- While some conversations need to be held among museums and libraries, others should include technologists and commercial organizations, such as the MIT Media Lab and IBM, as well as foundations.

Coordinate common endeavors

- There should be an easily accessible inventory of ongoing cooperative digital efforts.

- Cultural institutions' digitization efforts should be scaled up substantially. This increase cannot be achieved without an unprecedented level of interinstitutional cooperation and collaboration. The following are areas proposed for action:

  Archiving: A common repository or series repositories for digital masters, a "digital Fort Knox," to provide economies of scale and standardization for maintaining the digital assets of small, medium, and large institutions.

  Digitizing: A centralized digitizing service for cultural institutions.

  Portal site: A common entry site (e.g., "culture.org" or "content.edu") that will attract a large segment of Web users, be easy to navigate, and deliver several levels of access.

  Selection: Pilot projects in which institutions collaborate on digitizing related collections in different media and formats.

Gather more information

- There is a great need for cost surveys and studies from both museums and libraries.

- There is a need to further develop business models that include cost recovery and other self-sustaining features. Different strategies employed by museums and libraries, either individually or as members of consortia, must be critically evaluated and compared. Existing business models such as AMICO and JSTOR should be studied.

- User surveys and studies are critical in order to know who uses a site, what their expectations are, and how well they are being met. The Web survey done by CLIR is only a beginning.

- Focus groups of various users, from scholars to students to casual users, should be conducted.

Address intellectual property issues

- Libraries and museums must "keep counsel close at hand" when making selection decisions.

- The cultural community should decide how to influence decisions such as whether the use of thumbnail rendering of images constitutes fair use.

- Cooperative licensing agreements should be developed and agreements should be standardized.

APPENDIX I

Conference Participants

Rick J. Ashton
City Librarian
The Denver Public Library

Scott Bennett
University Librarian
Yale University

Francis X. Blouin
Director
Bentley Historical Library
The University of Michigan

Laura Campbell
Director
National Digital Library Program
Library of Congress

Stanley Chodorow
Special Associate of the President for Instructional Technology
University of California

Spencer Crew
Director
National Museum of American History

Jeff Field
Acting Director
Division of Preservation and Access
National Endowment for the Humanities

Barbara Franco
Executive Director
Historical Society of Washington

Douglas Greenberg
President and Director
Chicago Historical Society

José-Marie Griffiths
University CIO and Executive Director
Information Technology Division
The University of Michigan

Brian Hawkins
President
EDUCAUSE

Peter B. Hirtle
Manager, Digital Access
Co-Director, Cornell Institute for Digital Collections
Cornell University

Stanley N. Katz
Director
Center for Arts and Cultural Policy Studies
Woodrow Wilson School
Princeton University

Anne Kenney
Associate Director, Dept. of Preservation and Conservation
Co-Director, Cornell Institute for Digital Collections
Cornell University

Leila Kinney
Assistant Professor
Department of Architecture
Massachusetts Institute of Technology

Heike Kordish
Deputy Director, Research Libraries
The New York Public Library

Clifford Lynch
Executive Director
Coalition for Networked Information

Deanna Marcum
President
Council on Library and Information Resources

Mary Minow
Attorney and Adjunct Professor of Library Science
San Jose State University

Alan B. Newman
Executive Director of Imaging
The Art Institute of Chicago

Doralynn Pines
Associate Director for Administration
Metropolitan Museum of Art

Joyce Ray
Director, Office of Library Services
Institute of Museum and Library Services

Bernard Reilly
Director of Research and Access
Chicago Historical Society

Christine Ruotolo
Associate Director
Electronic Text Center, Alderman Library
University of Virginia

Samuel Sachs, II
Director
The Frick Collection

Patterson Sims
Deputy Director of Education and Research Support
Museum of Modern Art

Abby Smith
Director of Programs
Council on Library and Information Resources

Carl Smith
Department of English
Northwestern University

Katherine Spiess
Asst. Director for Strategic Initiatives
National Museum of American History

Christine Steiner
Attorney-at-law

APPENDIX II

Assessing Institutional Web Sites

Summary of Report

Prepared by John Chadwick, John H. Falk, and Brigitte O'Ryan
for the Council on Library and Information Resources

The Council on Library and Information Resources engaged the Institute for Learning Innovation to conduct a pilot study on museum and library Web sites. The purpose of the research was to begin to understand why and for whom institutions develop Web sites and to gain a better understanding of the needs of online users. The Institute for Learning Innovation developed the following research questions:

- Why are institutions building Web sites?
- Are the goals and objectives clearly stated and written down?
- Do all parties within an institution share the same goals and objectives for the site?
- Who does each institution assume uses its Web site?
- Who actually uses the sites?
- Why does the institution think Web users visit the site?
- Why do actual users visit the site and how do they find it?
- What do developers think site users take away from the experience?
- What do site visitors actually take away from the experience?

Methods

The Council on Library and Information Resources selected three libraries and three museums for this study, based on their participation in the conference *Collections, Content, and the Web*, held October 5-7, 1999, at the Chicago Historical Society. The six institutions were Yale University Library, Cornell University Library, New York Public Library, the Chicago Historical Society, the Art Institute of Chicago, and the Frick Collection.

Three study methods were used:

1. site visits were made to the participating institutions and interviews with held with key personnel,
2. a survey was sent to the participating institutions to solicit feedback from key administrators about the creation and management of their Web sites, and
3. an online user survey was linked to each institution's Web site, with results relayed directly to ILI's Web site.

Results and Discussion

Two types of organizational structures were represented in this study: (1) academic libraries, which use a decentralized computing system that is consistent with the organization of their campus libraries, and (2) public institutions, including the New York Public Library and the three museums, which use a centralized computing system. The libraries at Yale University are linked with individual academic units at the university. This traditional affiliation is reflected in Yale's Web site, which makes available online more than 40 libraries and special collections. The Cornell University Library system comprises more than 20 libraries and special collections. The New York Public Library, the largest public library system in the world, has a centralized management system for its Web-based projects. The Chicago Historical Society, the Art Institute of Chicago, and the Frick Collection are private institutions with facilities of varying sizes. Their collections also vary greatly in size and scope. Among the three museums, the departments responsible for managing online resources differ. It is important to bear in mind these differences when considering the results of this research project.

Why are institutions building Web sites?

All six institutions indicated that the definition of success for collections-based institutions is shifting from the size of collections to the services they offer. All the institutions view the Web as critical in redefining institutional success and helping to meet the new definition of success. Many of the institutions also said the Web would help increase their visibility. Some museums hoped the Web would increase visits to the museums.

Are the goals and objectives clearly stated and written down?

All the institutions participating in this research project indicated that they developed goals and objectives for their Web sites; however, only four sites had a set of written goals that had been distributed to staff members. Many respondents indicated that the goals had been shared informally through e-mail messages and face-to-face meetings. Within an institution, responses seemed to reflect confusion over what may or may not be goals and objectives for the Web site. For example, some respondents at one site indicated that there were indeed goals, while another respondent from the same institu-

tion stated that there were no written goals. The academic libraries have a much more difficult challenge in developing goals because so many libraries and special collections are closely linked to academic units. What may be appropriate goals for the main library may not be so for a special collection. Ultimately, each institution will need to develop goals and objectives that address its unique situation.

Do all parties within an institution share the same goals and objectives for the site?

If an organization has written goals and objectives, one might assume that everyone in the organization would know about and share them. The research results indicate that this assumption is not valid. The survey data show that, even when goals and objectives exist, not all organizations are communicating them. Organizations should be aware of the need to communicate larger goals and objectives to the staff involved in developing and maintaining online resources. If the goals and objectives of a Web site are not clear, the site may not be properly organized, and the online visitor may not be able to fully use the resources of the site.

Who does each institution assume uses its Web site?

Museums most often indicated the general public as their primary audience, while academic libraries stated they developed their Web sites to reach students, faculty, and staff. It is important to understand that users of both museum and library Web sites cannot always be classified into neat groups such as students, faculty, and staff. The Web has given these organizations an audience beyond the walls and traditional users of their institutions. For example, Cornell University maintains the United States Department of Agriculture (USDA) Economics and Statistics database. The reports of this database are intended to reach a global audience. Although Cornell expected that the Web site would serve primarily faculty, students, and staff, the concept of Cornell's audience changed after the Web site was established. The largest growing community of Web site users consists of practitioners in the field.

Who actually uses the sites?

More women (62 percent) than men (38 percent) completed the online survey. The mean age of participants was 39.4 years. The mean age for men completing the survey was slightly higher than that for women. The youngest person to complete a survey was 12 years old; the oldest was 86 years old. The mean age of those who completed a survey at a museum Web site was slightly higher than that of respondents who completed the survey at a library site.

Almost half of the respondents said the Web site would have no impact on how often they visited the physical site. About 33 percent of the respondents indicated that the Web site influenced them to visit the physical location more often, and 20 percent said the Web site influenced them to visit less often. Most of the respondents who claimed they would visit the institution more often were museum

site visitors, while most who indicated they would visit the institution less often were academic library site visitors. The Web master for one university's medical school library echoed this finding, noting that the use of the Web has increased while the number of people coming to the medical school library has decreased. This is because the Web has increased the library's ability to provide time-critical information and research to its intended audience.

Although two out of five respondents from the academic libraries indicated that they were visiting the physical site less often, an administrator at one site said that the availability of full-text journals online has actually increased the number of people coming to the library and asking for a hard copy. Another respondent noted that the Web site enabled her to do research even when the library is closed, and the time she does spend in the library is now much more efficient and productive.

Respondents to the user survey also indicated they visited the Web site quite frequently, with 55.4 percent of online visitors to museum Web sites and 70 percent of library users visiting Web sites at least once a month. None of the museum respondents indicated that they were first-time visitors to the Web site while libraries had a substantial number of first-time users.

See tables 1–3 and figure 2.

Why does the institution think Web users visit the site?

Although both libraries and museums cited information as a primary reason people visit their Web sites, the types of information and the scope of the information varied as a function of the mission and audience of the institutions. Most library administrators said that Web users visit their sites for convenient access to information about resources and collections. Several administrators also noted that users were motivated to use the Web site because of its convenience in providing electronic access to the institution 24 hours a day, seven days a week. Some museum administrators also listed education and entertainment as reasons for using the site.

Why do actual users visit the site and how do they find it?

People visit libraries for different reasons than they visit museums; therefore, it is not surprising that the visitors to library and museum Web sites have different agendas. Virtually half of the online visitors to the libraries indicated that they were visiting the site to conduct academic research; more than a third indicated that they were visiting to conduct personal research. Less than 15 percent visited for personal growth or to plan a visit to the library. Conversely, for museum site users, the reason most often cited for the visit was to seek information required for planning a visit to the physical site; this was followed closely by personal research and personal growth. Academic research was cited as a reason by only a fifth of the respondents visiting museum sites, even though the three museums in the study have strong research libraries associated with their institutions. None of

the museum administrators mentioned planning an actual visit as a reason for a visit to the Web site.

Visitors to a museum site were more likely to find the site with a search engine, while visitors to a library site were more likely to enter the site either by typing the site address or through the default Web page on the browser. The lowest percentage of online visitors found the site through links at other sites. Museums and tourism agencies often provide links to museum Web sites. These links are critical to maintain in order to reach a broad audience, but they accounted for the lowest percentage of the total online visitors.

See figures 3–7 and tables 4, 6, 7.

What do developers think site users take away from the experience?

Most of the Web administrators were vague in answering this question. Generally, administrators hoped that visitors were able to find the information they were looking for when they came to the sites, such as information about the collections and resources, and about the institution. Some museum administrators said they hoped Web users had an educational experience and were entertained.

What do site users actually take away from the experience?

More than 90 percent of the respondents rated the information on the library and museum Web sites as either "very reliable" or "somewhat reliable," and they return to sites regularly. Ease of access to information was also rated highly among respondents from the library Web sites. Overall satisfaction with the Web sites was quite high; 84 percent of the respondents were either "somewhat satisfied" or "satisfied" with their online experience.

See figures 4, 8, 9, 10, 11 and tables 5, 8.

Conclusion

This pilot study has provided some insights into the perceptions of those who develop library and museum Web sites and those who use those sites. We have also created a model that can be used by other institutions in gaining a better understanding of their online visitors.

We have found points of convergence between institutional intent and user needs as well as points where the two diverge. Although the cultures of museums and libraries developed from very different roots and demands, the political and cultural realities of the twenty-first century are fostering a convergence in both mission and practice. This convergence is nowhere more evident than in the use and application of technology. Although significant differences exist, the similarities are becoming more evident and important.

Research in all facets of Web use is required, but further work is particularly recommended on several issues that emerged from this project. The link between visiting an institutional Web site and making a real visit needs to be explored. Understanding this connection can help libraries and museums create stronger connections for their

visitors between the physical and virtual sites. For museums, a well-designed virtual experience may lead to more visits to the physical site, which can lead to increased revenue. Libraries will be able to make better and more informed budgetary decisions that will lead to improved services for visitors to the physical and virtual sites.

In addition, the frequency of visits to the Web sites in this study deserves further inquiry. By understanding the needs of repeat visitors to library and museum Web sites, administrators and developers can take advantage of the technology to better serve an audience that may not be able to visit the physical site. Finally, this research points up the need for greater clarity of goals and objectives and tools that are more refined to assess both costs and benefits of this new information tool.

Appendix II Tables and Figures

Table 1. Mean age of respondents to user survey by type of institution

Type	Mean	Std. Deviation	Low	High
Library	38.8	14.8	14	86
Museum	40.0	14.0	12	75

Table 2. Mean age of all respondents

Gender	Mean	Std. Deviation	Min.	Max.
Women	38.3	13.5	12	74
Men	40.7	15.5	15	86

Table 3. Gender of respondents by type of institution

Type	Female	Male
Library	54.3% 133	45.7% 67
Museum	71.0% 164	29.0% 112
Total	62.4% 297	37.6% 179

Table 4. Why online visitors are coming to the Web site by type of organization

Type	Academic research	Personal research	Personal growth	Plan a visit to the library	Plan a visit to the museum
Library	49.8% 122	36.7% 90	8.2% 20	5.3% 13	NA
Museum	20.3% 47	26.4% 61	25.5% 59	NA	27.7% 64
Total	35.5% 169	31.7% 151	16.6% 79	2.7% 13	13.4% 64

BEST COPY AVAILABLE

Table 5. What is important about the information at this Web site by type of institution

Type	Reputation of organization	Depth of information	Easy to access information	Reliability of information
Library	7.1% 17	18.8% 45	45.4% 109	28.8% 69
Museum	3.5% 5	18.2% 26	32.9% 47	45.5% 65
Total	5.7% 22	18.5% 71	40.7% 156	35.0% 134

Table 6. Frequency of visits to Web site by type of organization

Frequency of visit	Library	Museum
First visit	22.4% 55	0
Daily	25.3% 62	18.6% 43
Weekly	31.6% 78	12.6% 29
Monthly	12.7% 31	24.2% 56
Seldom	6.9% 17	37.2% 86
Yearly	0.8% 2	17.4% 17
Total	245	231

Table 7. Where respondents were physically located by type of institution

Type	Offsite	Onsite
Library	77.6% 190	22.4% 55
Museum	80.5% 186	19.5% 45
Total	79.0% 376	21.0% 100

BEST COPY AVAILABLE

Table 8. Meeting expectations by type of institution

Type	Failed to meet expectations	Met some expectations	Met most expectations	Met or exceeded expectations
Library	3.0% 7	15.1% 35	47.0% 109	34.9% 81
Museum	20.0% 42	14.8% 31	33.3% 70	31.9% 67
Total	11.1% 49	14.9% 66	40.5% 179	33.5% 148

Figure 1. Number of responses from each of the participating institutions

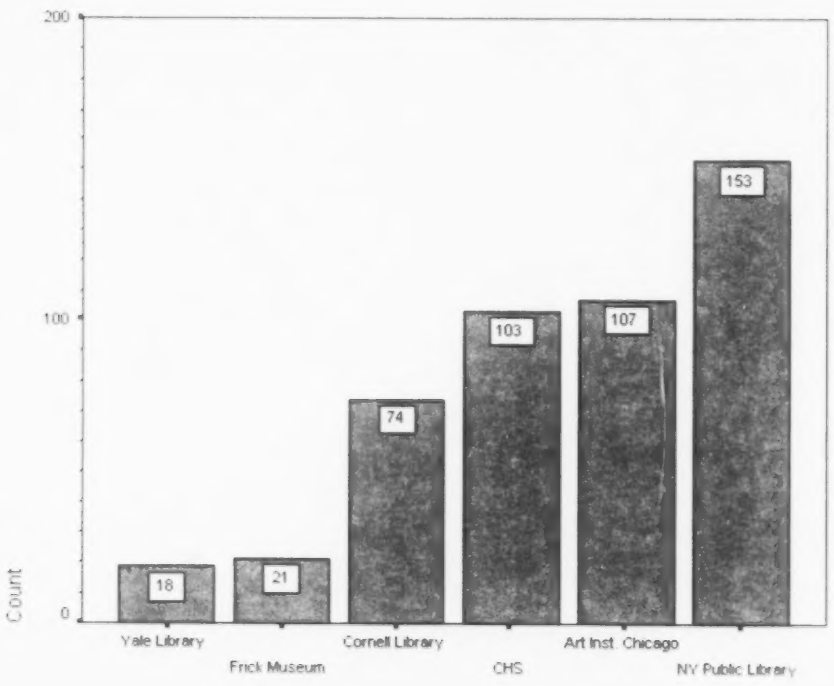


Figure 2. Percent of respondents by gender

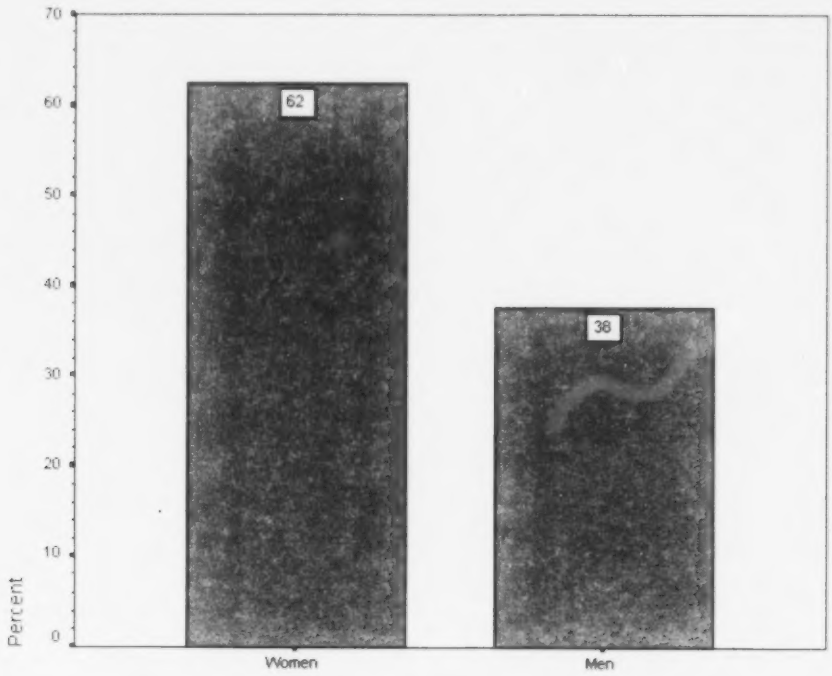


Figure 3. Why are you visiting this Web site today?

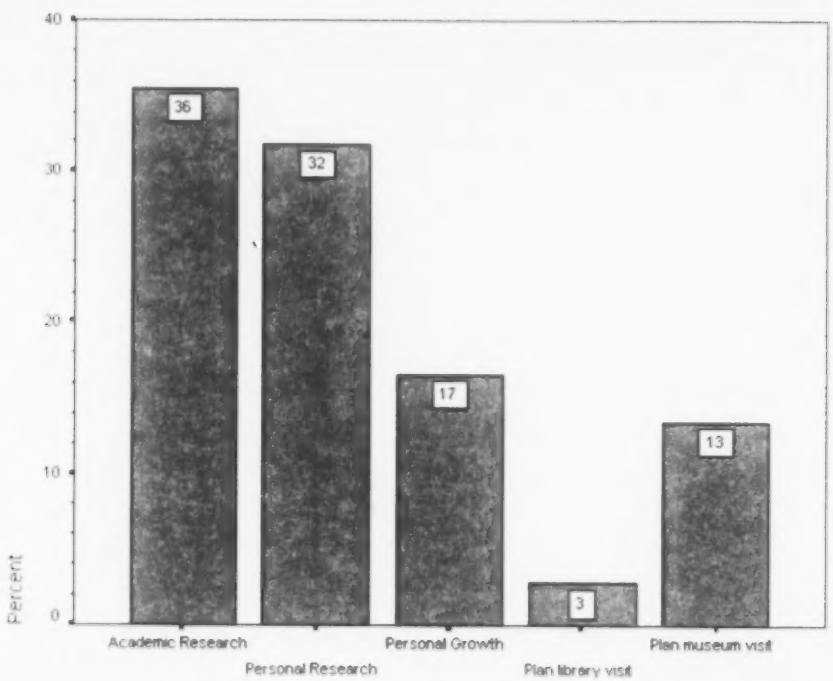


Figure 4. What is most important about the information at this Web site?

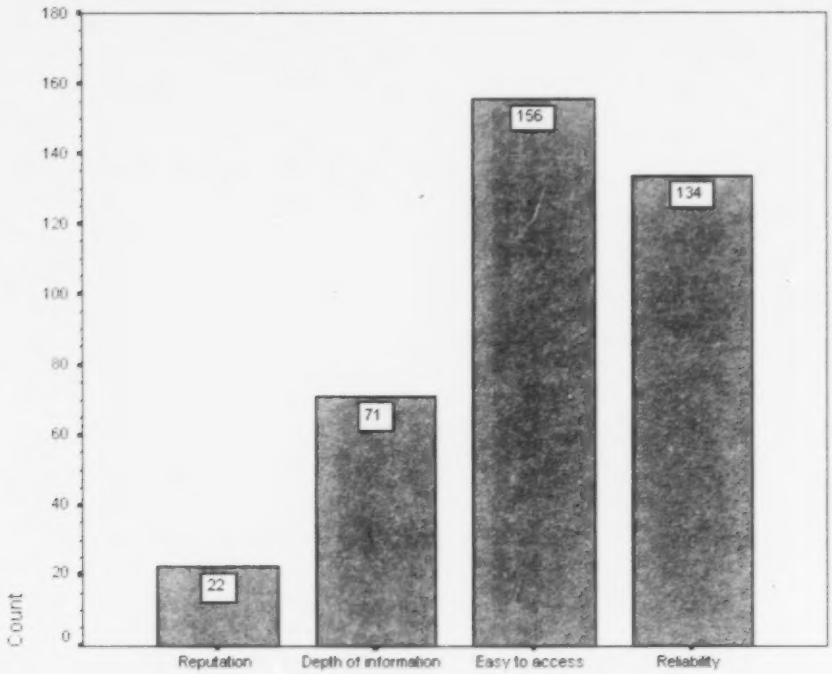


Figure 5. Total number of respondents who indicated how likely they would visit the physical site

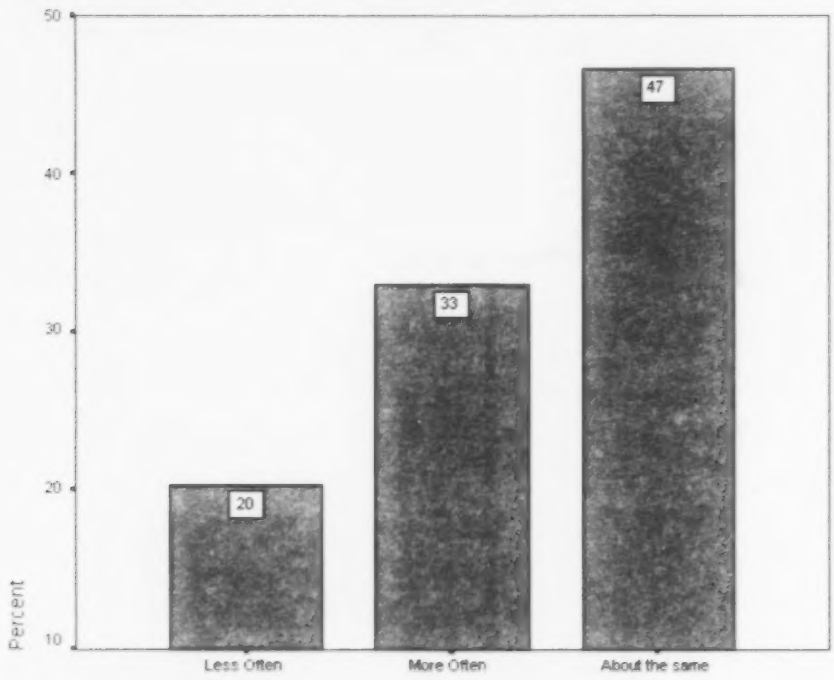


Figure 6. Where respondents said they were physically located relative to the site they were visiting

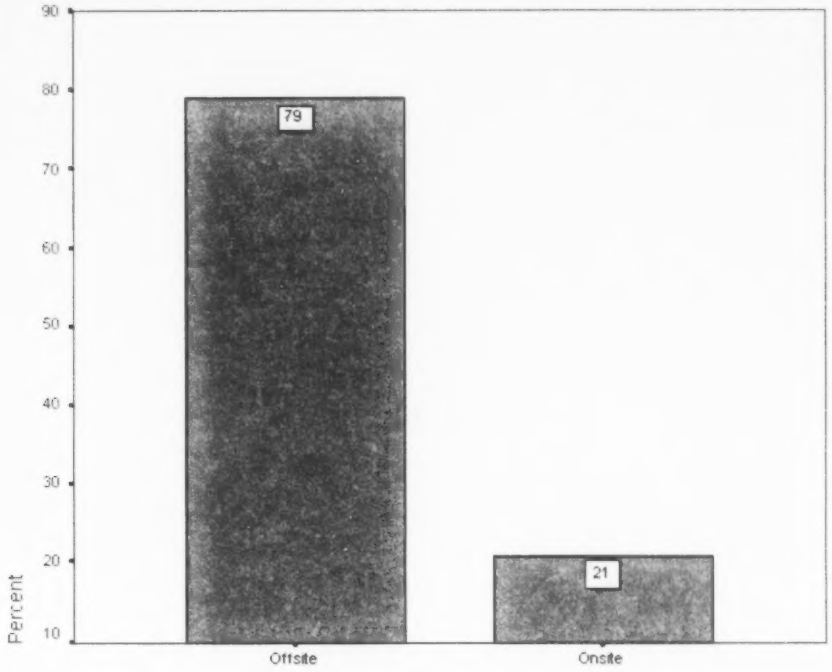


Figure 7. How online visitors found the Web site

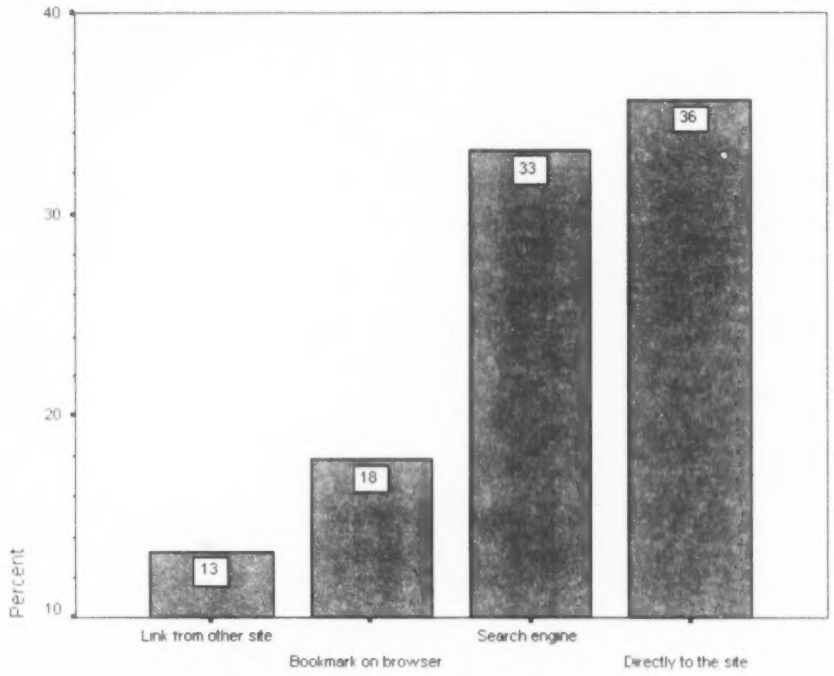


Figure 8. Aggregate of responses of users rating the reliability and trustworthiness of the sites they were visiting

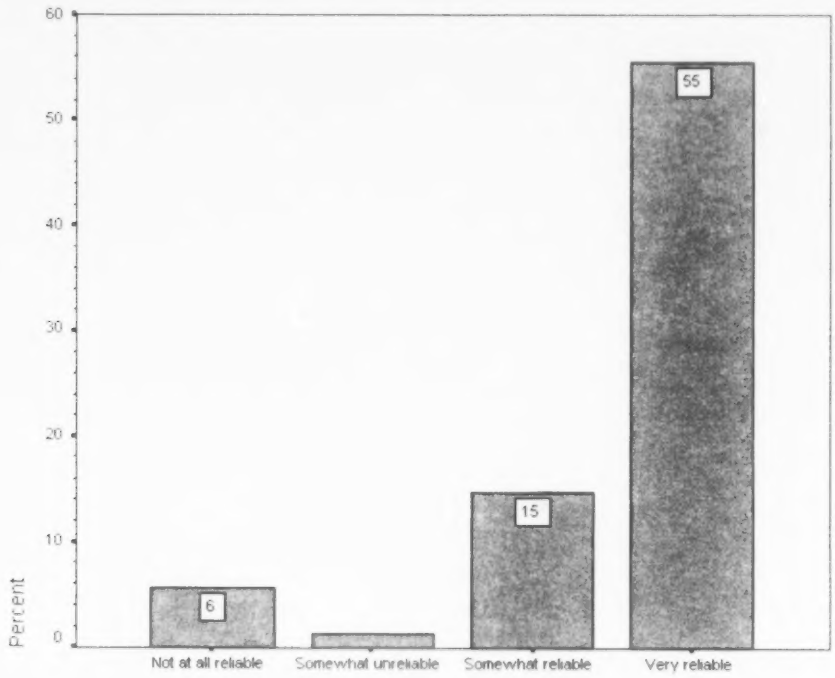


Figure 9. Ranking of importance for e-mail contact by all respondents

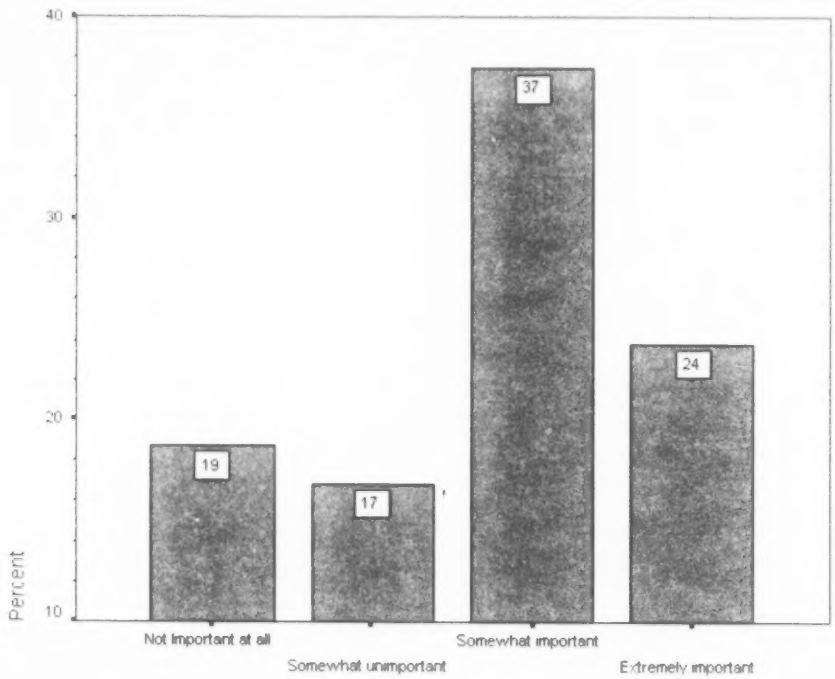


Figure 10. Overall user satisfaction with Web sites

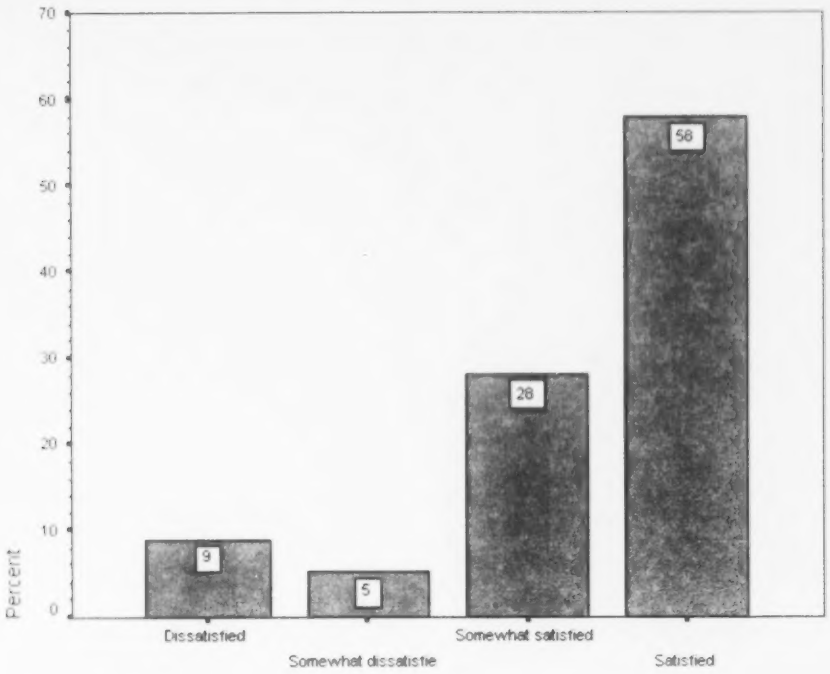
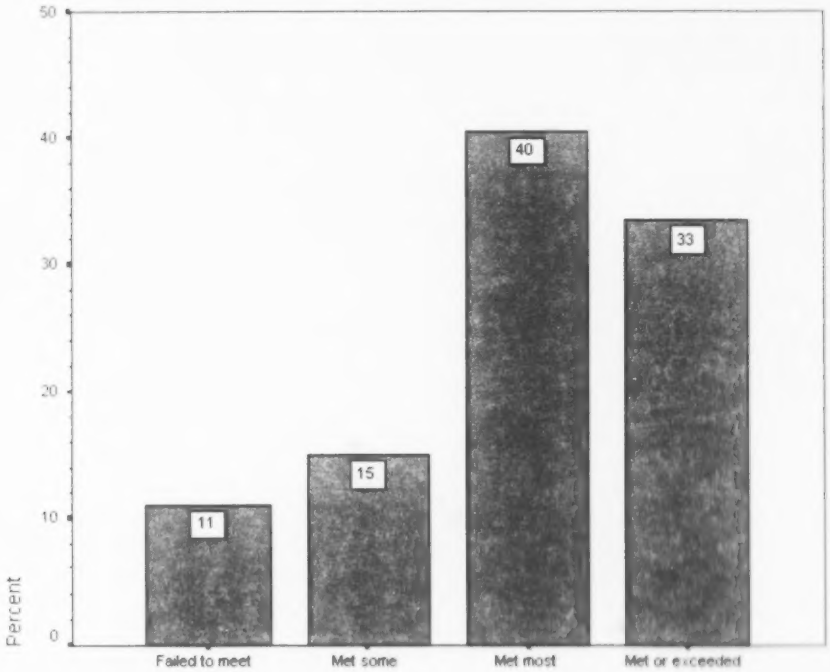


Figure 11. Overall user rating of Web sites meeting expectations



BEST COPY AVAILABLE

COUNCIL ON LIBRARY AND INFORMATION RESOURCES

1755 MASSACHUSETTS AVENUE, NW, SUITE 500, WASHINGTON, DC 20036-2124
Telephone: 202-939-4750 • Fax: 202-939-4765 • E-mail: info@clir.org • Web: www.clir.org